A TEXTBOOK OF PHYSICS

BLACKIE & SON LIMITED
50 Old Bailey, LONDON
17 Stanhope Street, GLASGOW

BLACKIE & SON (INDIA) LIMITED
Warwick House, Fort Street, BOMBAY

BLACKIE & SON (CANADA) LIMITED
TORONTO

A TEXTBOOK OF PHYSICS

BY

E. GRIMSEHL

EDITED BY

R. TOMASCHEK, D.Phil.
Professor of Physics, the University of Marburg

Authorized Translation from the Seventh German Edition by

WINIFRED M. DEANS, M.A.(Cantab.), B.Sc.(Aberd.)

VOL. IV
OPTICS

BLACKIE & SON LIMITED
LONDON AND GLASGOW

BOOK
PRODUCTION
WAR ECONOMY
STANDARD

THE PAPER AND BINDING OF THIS BOOK CONFORM TO THE AUTHORIZED ECONOMY STANDARDS

First published 1933
Reprinted 1941

Printed in Great Britain by Blackie & Son, Ltd., Glasgow

FROM THE PREFACE TO THE SEVENTH GERMAN EDITION

The present volume contains those parts of Optics which deal with the *propagation* of light; information relating to the *emission* of light, which, indeed, cannot be explained by any theory of a continuum, will be found in Vol. V. The subjects treated here follow naturally on the discussion of electromagnetic waves in Vol. III. In the main it has been possible to utilize the material of the previous edition; but the fact that electromagnetic waves had already been studied enabled me to adopt a somewhat different order of treatment, and also to lay greater emphasis throughout on the wave theory, which in the light of long teaching experience I consider very necessary. In many places the existing material has been amplified or new matter added, e.g. in the passages dealing with photometry, dispersion, the wave theory of image formation, the optical properties of X-rays, the resolving power of spectroscopes, the Kerr effect, and methods for determining the velocity of light.

R. TOMASCHEK.

MARBURG ON THE LAHN,
October, 1931.

PREFACE TO THE ENGLISH EDITION

I have taken the opportunity of the publication of this edition to make a few corrections and improvements. With a view to meeting the needs of English readers, some new matter has been added. I wish to thank the translator for the adequate manner in which she has carried out her task, and hope that in its new dress the work will greatly extend the circle of its friends.

R. TOMASCHEK.

MARBURG,
10th August, 1933.

CONTENTS

CHAPTER I

THE PROPAGATION OF LIGHT

CHAPTER II

THE INTENSITY OF LIGHT

CHAPTER III

GEOMETRICAL OPTICS: REFLECTION AND REFRACTION

CHAPTER IV

GEOMETRICAL OPTICS: LENSES

CHAPTER V

GEOMETRICAL OPTICS: ABERRATION

CHAPTER VI

GEOMETRICAL OPTICS: APPLICATIONS

CHAPTER VII

THE INFLUENCE OF WAVE-LENGTH ON REFRACTION PHENOMENA

CHAPTER VIII

THE INTERFERENCE OF LIGHT WAVES

CHAPTER IX

THE VELOCITY OF LIGHT

CHAPTER X

THE POLARIZATION OF LIGHT

CHAPTER XI

OPTICAL PHENOMENA IN THE ATMOSPHERE

CHAPTER XII

PHYSIOLOGICAL OPTICS

COLOURED PLATE

OPTICS

CHAPTER I

The Propagation of Light

1. General Aspects of the Propagation of Light.

Introduction.—In the volume on Electricity (Vol. III) we met with light as an electromagnetic wave motion originating in the molecules and atoms themselves. In what follows we shall consider the propagation of light in more detail. Although in general we shall speak of visible light only, i.e. waves in the region 800–400 $m\mu$, most of the relationships we shall mention apply with very great accuracy to the infra-red and ultra-violet regions, i.e. to waves from about 10μ to $100m\mu$ (Vol. III, p. 646, fig. 47). Shorter waves exhibit peculiarities which depend partly on their high frequency and partly on the electrical structure of matter. Many of the laws which hold for visible light, however, are true for all electromagnetic waves. Indications of the range of validity of the facts described will therefore be given where necessary.

Source of Light. Light Quantum. Unpolarized Light.—A body which itself emits light (e.g. the sun, the fixed stars, an electric lamp) is called a **luminous body** or **source of light.** We must think of a source of light as consisting of an immense number of emitters sending out trains of electromagnetic waves more or less simultaneously and independently of one another. Any one such emitter we may call a centre of homogeneous radiation. At present we still do not know the nature of the wave-trains starting from the individual centres, whether they are damped and if so to what extent, whether they are sent out in definite favoured directions (Vol. III, p. 626, fig. 27) or are radiated with equal intensity in all directions. (For further details see Volume V.)

For the phenomena to be discussed in the present volume only the following properties, which we shall mention here but not establish till later, are of importance. (1) Each one of these wave-trains keeps

together; it is **coherent** or **continuous** (Vol. III, p. 641), i.e. the waves follow one another at the same distance and in the correct phase over a length of several metres. As is shown by interference phenomena (see e.g. p. 15) the individual wave-train is **polarized** (Vol. III, p. 638). As we found in the case of electromagnetic waves, the plane in which the electric vector vibrates is the same for the whole wave-train (e.g. the plane of the paper in fig. 1*a*). The magnetic vector vibrates in the plane at right angles to the electric vector and to the direction of propagation. A single wave-train of this kind is called a **light quantum**. Its length amounts to several metres. As the velocity of propagation of light in vacuo is 3.10^{10} cm./sec. (Vol. III, pp. 195, 196) each wave-train (reckoned from a fixed point past which the train of waves moves) persists for about 10^{-8} sec. (2) In general the phase and direction of polarization of the wave-trains which are emitted in enormous numbers by a source of light vary quite arbitrarily from wave-train to wave-train; the wave-trains are mutually **incoherent**.

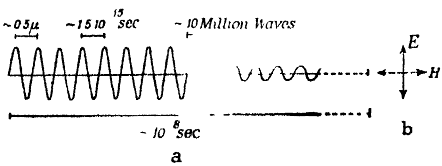


Fig. 1*a*, *b*.—Single train of light waves. (In *b* the direction of propagation is at right angles to the plane of the paper.)

Ordinary unpolarized light, then, consists of this type of wave motion, starting from the source of light and propagated in a homogeneous medium, e.g. a vacuum or air, in straight lines in all directions. The light radiated by a source accordingly consists of an immense number of individual flashes of light; in Vol. V we shall learn how their number may be calculated. Thus when we look at a definite point of a source of light, our eye is excited without intermission by separate trains of waves, each with its own phase and direction of polarization, arriving quite irregularly, often on top of one another, but sometimes with gaps between. It is important that the student should master this conception of the nature of the light radiated by a visible source. Further details will be discussed and established later.

In what follows we shall often be able to neglect the dimensions of the source of light in comparison with the other distances involved. We then speak of a **point source**. The distinction between a point source and a centre of homogeneous radiation is that the radiation from the former consists of a large number of mutually incoherent trains of waves.

Light Rays.—We shall now meet with another conception whose use greatly simplifies a large number of discussions.

We call the direction in which the light energy flows a ray of light. According to Poynting's theorem (Vol. III, p. 618) the ray of

light is at right angles to the plane of the electric and magnetic intensities, and forms the normal to the wave surface (Vol. II, p. 239). In a homogeneous medium the rays of light from a source at rest are straight lines (fig. 2): that is, the individual wave-trains are spherical waves (or at least parts of spherical waves) with the source as centre.

In ordinary speech we apply the term ray (or beam) of light to a more or less extended rectilinear strip of light such as arises when sunlight comes in through a small hole in an opaque screen. In physics we shall meanwhile neglect the breadth of the ray or beam, i.e. we assume that it has the properties of a geometrical straight line.

A large number of parallel rays of light considered together (to which we do ascribe extension in space) form what is called a **bundle** or **pencil of parallel rays.** A large number of rays of light diverging from one point form **a divergent pencil of rays.** The pencil of parallel rays is associated with a plane wave surface, the divergent pencil with a curved wave surface.

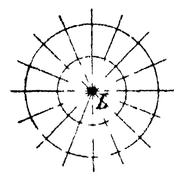


Fig. 2.—Radiation from a very small source emitting light in all directions (wave surfaces and rays)

Reflection and Refraction.—If light waves fall on a body they (as in the case of any wave motion that encounters a medium with different properties) partly are reflected and partly enter the medium and are thereby refracted (Vol. II, p. 248).

It is the light thrown back by bodies on which light falls that makes the bodies visible. Such bodies are said to be **non-self-luminous,** in contradistinction to sources of light, which are **self-luminous.** Between these two there is an important difference. If we look at an extended source of light, say the glowing filament of an electric lamp, the wave-trains which reach the eye at any one instant from two different points of the filament are **incoherent,** i.e. their phases and directions of polarization are quite different. If, however, we look at the filament by daylight with the current off, the wave-trains reaching the eye from the same parts of the filament are coherent, for the light at any instant now arises from the reflected parts of one and the same wave which has come from a source of light and has been reflected at the filament.

The non-reflected part of the light enters the body and either passes through it or is absorbed in it, i.e. its energy is changed into another form, usually into heat. Bodies through which light passes are said to be **transparent** or **translucent**: bodies which let no light pass through them are called **opaque.**

There is no definite boundary between transparent bodies and opaque bodies; on the contrary, bodies which in ordinary usage are described as transparent, such as clean water, are opaque in thick layers (complete darkness reigns in the depths of the ocean). On the other hand, opaque bodies allow the passage of light

when in sufficiently thin layers (thin gold leaf appears green by transmitted light, silver leaf blue). In general light is partly reflected by bodies, partly transmitted, and partly absorbed. A body which absorbs practically all the light is opaque and appears black by reflected light. A body which absorbs all the light falling on it is said to be a *perfectly black body*: such a body does not actually exist, the "perfectly black body" being an ideal conception. If we speak of an opaque body, we mean a body which lets through so little of the light that the transmitted light cannot be detected by the senses or by other means. We must restrict the idea of transparency (in the wider sense of transparency to any electromagnetic waves) still further in view of the fact that a body may be transparent to one wave-length but opaque to another. Thus, for example, ordinary window-glass does not transmit the health-giving ultra-violet rays of sunlight; it is opaque to them (p. 161). On the other hand, black paper transmits infra-red rays (p. 162). The tissues of the body are opaque in thick layers, but transparent to X-rays. Even in the visible region, however, the same phenomenon may occur. A body may be transparent to green light, opaque to red light. It then appears coloured. These special cases will be discussed later (p. 156).

White and Coloured Light.—The radiation from a source of light appears *white* if all the frequencies corresponding to visible light occur in it with roughly the same intensity, i.e. if on the average the number of individual wave-trains is the same for each frequency.

The totality of frequencies emitted by a source of light is called the spectrum of the source.

If some parts of the spectrum are wanting, i.e. if wave-trains of a definite frequency range are absent from the radiation or only present in small quantities, the radiation appears more or less coloured, and the deeper in tint, i.e. the less mixed with white, the more extensive the spectral region which is wanting. Radiation of a single definite wave-length is called **monochromatic radiation**; this gives the sensation of a pure colour, the individual wave-lengths corresponding to the colours given on p. 646 of Vol. III. Experiments dealing with these points are discussed in § 1 of Chapter VII (p. 150).

2. Situation of the Source of Light.

If we bring an opaque body between a source of light and the eye, we cease to see the source of light when the source of light, the opaque body, and the eye lie in one straight line. It follows from this that (in ordinary circumstances) light is propagated in straight lines. The knowledge that the path of the ray of light is a straight line enables us to determine the *direction* in which a source of light or illuminated body lies. For psychological reasons we always look for the starting-point of the light in the backward prolongation of the rays of light entering the eye; by observing a source of light from *one* point only and with one eye, however, we cannot tell at what *distance* it lies from the eye.

Suppose a source L is emitting rays of light (fig. 3), two of which are represented in the figure by LC_1 and LC_2; if our eye is at C_1 we can

only say that the source of light lies in the *direction* C_1L. Whether it is situated at M_1 or L or N_1 we cannot tell. Similarly, if our eye is at C_2, we can only say that the source of light lies in the direction C_2L. If we observe with this eye alone we may equally well expect the source of light to be at M_2 or N_2.

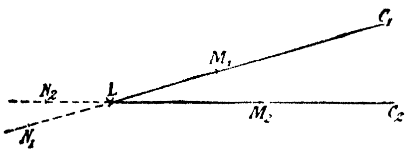


Fig. 3.—Determination of the position of a source of light

If, however, we observe with *both* eyes simultaneously, i.e. look along the ray C_1L with one eye and along the ray C_2L with the other, or if we observe the source of light from C_1 and C_2 in rapid succession, we also obtain the *distance* of L, its position being determined by the intersection of the two rays. The position of L is the more accurately determined the nearer the angle C_1LC_2 approaches a right angle. Hence in order to estimate the distance of a remote object accurately, we observe it in rapid succession from two points wide apart, by moving from one point to the other. For near objects observation with the two eyes simultaneously will suffice.

In the middle of the room we hang a ring from the ceiling so that the plane of the ring passes through one of our eyes, and approach the ring with the other eye shut. If we then try to put a pencil sideways through the ring we find we can only do so after many vain attempts. If, on the other hand, we keep both eyes open, we can do it the first time we try. In the first case the one eye gave us only one direction; in the second case the two eyes gave us two directions, and hence their intersection.

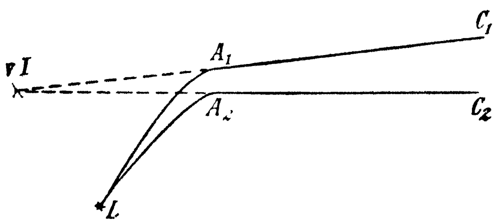


Fig. 4.—Virtual image of a source of light

If we know nothing more about the actual path of the rays emitted by a source of light, we shall be deceived about the direction and distance of the source if the rays for any reason fail to travel in straight lines.

If, for example, the directions of the light rays starting from L (fig. 4) are altered at A_1 and A_2 in such a way that they run in straight lines from A_1 to C_1 and from A_2 to C_2, the eyes of the observer at C_1 and C_2 look for the source in the backward prolongation of these two rays, i.e. at the point *v*I; it appears as if the source of light were at *v*I. In this case we call ***v*I the virtual image** of the source L.

If the directions of the light rays starting from L are altered at A_1 and A_2 (fig. 5) in such a way that before reaching the eye they intersect at *r*I and then proceed to C_1 and C_2, the eyes at C_1 and C_2 look for the

source at rI, the point of intersection of the rays A_1C_1 and A_2C_2; the source of light appears to be situated at this point. In this case we call rI the **real image** of the source L.

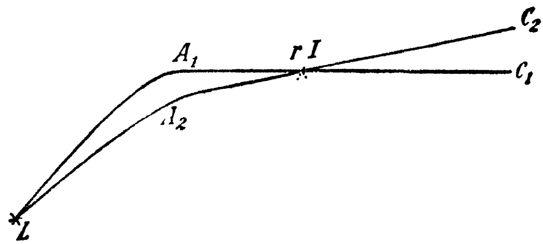


Fig. 5.—Real image of a source of light

A virtual image is the point of intersection of the prolongations of the rays, a real image the point of intersection of the rays themselves.

3. The Reflection of Light.

If a body is visible from all sides, it must be emitting light in all directions. A piece of paper on which the sun is shining is visible in *all* directions, because it sends back light in all directions, although the light falling on it consists of parallel rays of light all coming from the same direction (**diffuse reflection**). A flat glass plate illuminated by sunlight, on the other hand, appears so bright when looked at from a certain direction that we are dazzled by the reflected light, whereas from other directions it is almost invisible. The glass plate throws back the light falling on it in one definite direction only (**regular reflection**). All bodies which are extremely smooth (highly polished), especially metals, likewise reflect regularly. If a parallel pencil of rays passes through a hole in an opaque screen, it is not visible unless the beam enters the eye. If, however, dust or smoke is brought into the path of the beam, the path of the beam becomes visible from the side also, the individual particles in the dust or smoke scattering part of the light in all directions (p. 149).

If a beam of light which has been made visible by means of dust is allowed to fall on a flat glass plate or metal mirror, we find that the reflected ray of light has a perfectly definite direction, which depends on the direction of the incident light. If a straight line is drawn perpendicular to the plane of the mirror at the point where the light meets the mirror, the angle between the incident ray and this normal is called **the angle of incidence.** The angle between the normal and the reflected ray is called the **angle of reflection.**

The disc shown in fig. 6 can be rotated about an axis through its centre, to which a small plane mirror S is fixed. The edge of the disc is graduated, the 90° mark lying where SL, the normal to the mirror, intersects the circumference of the disc.

A ray of light incident on the mirror along the line LS is reflected in such a way that the reflected ray coincides with the incident ray.

If the disc is rotated so that the ray incident on the mirror in the direction ES makes the angle of incidence ESL with the normal SL, the ray is reflected in the direction SR, making the angle of reflection LSR with the normal.

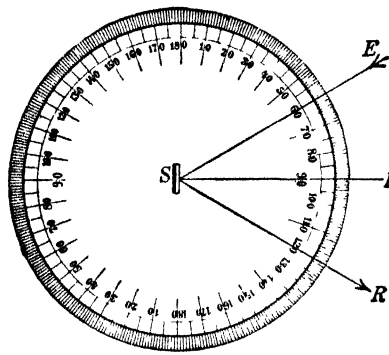


Fig. 6.—Apparatus for reflection experiments

From such observations we obtain the **law of reflection.***

(1) *The incident ray, the normal at the point of incidence, and the reflected ray all lie in one plane.*

(2) *The angle of incidence is equal to the angle of reflection.*

If the mirror S is rotated through any angle, the direction of the incident light remaining the same, the normal is rotated through the same angle, and the angle of incidence and the angle of reflection are both altered by this angle; relative to the incident ray the reflected ray is therefore rotated through an angle which is *twice* the angle through which the mirror is rotated (see the mirror and scale arrangement illustrated on p. 34).

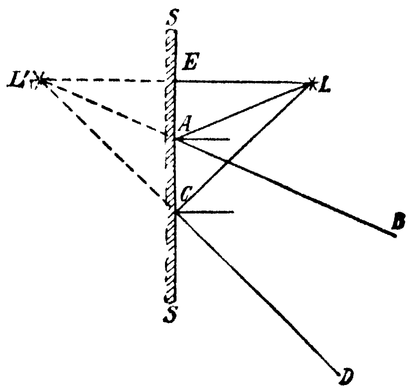


Fig. 7.—Reflection at a plane mirror

Plane Mirrors. — A point source of light L is situated in front of a plane mirror SS (fig. 7). The rays of light LA, LC are reflected in the directions AB, CD respectively. An eye on which a narrow pencil of rays with AB as central ray (p. 44) falls will expect the source of light to be in BA or BA produced. An eye on which a narrow pencil of rays with CD as central ray falls will expect the source of light to be in DC or DC produced. If we observe the two rays AB, CD simultaneously, one with each eye, the source of light appears to be situated at L′, the point of intersection of the two lines produced. The line joining L and L′ meets the surface of the mirror at E. From the law of reflection

* The discovery of the law of reflection is ascribed to Euclid of Alexandria (300 B.C.).

and the consequent congruency of the triangles behind and in front of the mirror it follows that:

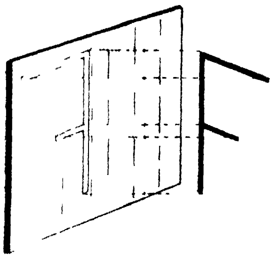


Fig. 8.—Object and image

LL′ is at right angles to the surface of the mirror and L′ is as far behind the mirror as L is in front of it.

L′ is the *virtual image* of L (p. 6), for it is the point of intersection of the *prolongations* of the rays of light incident on the eye. It is obvious at once that the rays *themselves* do not intersect at L′, as none of the rays under consideration are in the space behind the mirror at all.

The mirror image of an extended source of light is composed of the mirror images of the individual points of the source of light. From fig. 8 it follows that in reflection the up-down direction is unaltered, whereas the front-back direction is reversed. The right-left direction also appears reversed, to judge from the image, which is what one does when one looks at oneself in the mirror; from this point of view, therefore, the mirror image is laterally reversed.

4. The Propagation of Light.

The phenomenon illustrated in fig. 9 depends on the rectilinear propagation of light. AB is an extended source of light, e.g. a number of small lamps arranged to form the letter F, emitting light in all directions. At some distance is placed an opaque screen with a hole L. and at a somewhat greater distance the white screen S. Of the light rays which are emitted by A and form a divergent pencil of rays only a very small pencil of rays can pass through the hole L. This gives rise to a bright spot on the screen S. In the same way the perforated screen W separates out from all the pencils of rays one pencil which forms a bright spot on the screen S. These spots of light on the screen are so arranged as to form a figure which is similar to the object AB but is upside down, both left and right and up and down being interchanged (looking from AB). The figure on the screen is no proper image in the sense of § 2, p. 5: the smaller the opening, the sharper the figure, as then the circle in the figure corresponding to any point of the object is the smaller.

Fig. 9 illustrates the principle of the pin-hole camera * (fig. 10).

* Sometimes ascribed to GIAMBATTISTA DELLA PORTA of Naples (1538–1615), who describes the "camera obscura" with a lens. The pin-hole camera, however, was known long before his time; it is mentioned by LEVI BEN GERSON about 1321 as a means for observing the sun. FRANCISCUS MAUROLYCUS (1494–1575), of Constantinople and later of Messina, explained the round spots of light in the shadows of trees, produced by sunlight shining through the interstices of the leaves, on the basis of the rectilinear propagation of light. Even ARISTOTLE had observed that during partial eclipses these images of the sun are sickle-shaped.

The back wall of the camera is a piece of ground glass on which the reversed images of the objects in front of the hole may be observed.

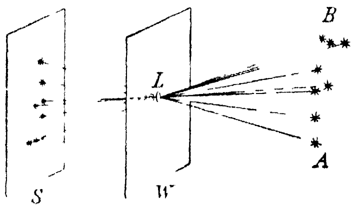


Fig. 9.—Principle of the pin-hole camera

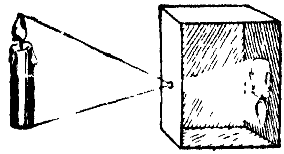


Fig. 10.—The pin-hole camera

This experiment reveals a new and wonderful fact. If we point the camera say at a sunny landscape, all its finest details may be noted on the ground glass. All the rays making up the varied picture have passed simultaneously through the tiny opening without the slightest disturbance. The individual wave motions have therefore been superposed on one another at the hole without affecting one another in the least, thus clearly showing that the superposition of electric and magnetic forces in the ether is strictly additive (i.e. according to the parallelogram of forces (Vol. I, p. 91)).

5. Shadows: Eclipses.

Shadows.—If we bring an opaque body K into the region illuminated by the point source of light P (fig. 11), a dark region or *shadow* S is produced behind the body.

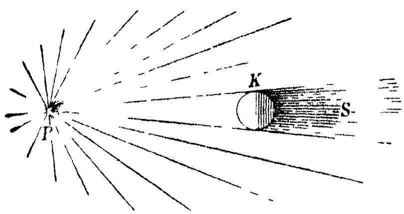


Fig. 11.—Umbra or region of total shadow produced by a point source of light

In a region illuminated by *two* point sources of light P and Q (fig. 12) the body K gives rise to three different shadow regions. In the region S there is no light at all (total shadow or **umbra**); the region H_1 is illuminated by P and not by Q; the region H_2 is illuminated by Q and not by P. The regions H_1 and H_2, which are only partially illuminated, form the **penumbra.**

In the same way the body K (fig. 13) illuminated by the three sources of light P, Q, R gives rise to the shadows S, H_1, and H_2. In the umbral region S no light is received; the penumbral regions H_1

and H_2 are illuminated by only one or two of the sources of light; the rest of the space is illuminated by all three sources of light.

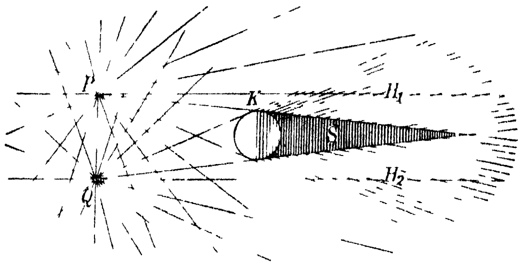


Fig. 12 —Umbra and penumbra produced by two point sources of light

If the source of light is an *extended surface* F (fig. 14), each point of it emits light rays, so that the opaque body K gives rise to the

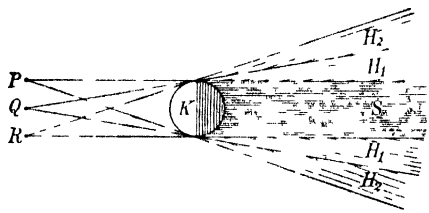


Fig. 13.—Shadows due to an extended source of light

umbra S and the penumbra H. In this case the umbra passes ***gradually*** into the penumbra and the latter into the fully illuminated region, while in figs. 11, 12, and 13 the regions of shadow are sharply bounded

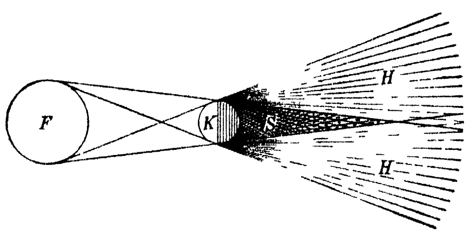


Fig. 14 —Shadow of a planet illuminated by the sun

off. Thus if the source of light used is *extended*, the shadows thrown by an opaque body (screen) are *blurred.*

Eclipses of the Sun and Moon.—The phenomenon illustrated in fig. 14 corresponds (apart from the relative sizes of the objects) to the shadows thrown during eclipses. If F represents the sun and K the earth, and the moon enters the shadow behind K, we have an *eclipse of the moon*. The eclipse is total if the

whole surface of the moon enters the shadow of the earth, partial if only part of the surface of the moon does so.

If F is the sun and K the moon, and the earth enters the shadow of the moon, we have an *eclipse of the sun*. An eclipse of the sun is total when the earth is in the umbra of the moon, partial when the earth is merely in the penumbra of the moon.

6. The Diffraction of Light.

Introductory Remarks.—If the openings and screens used in the experiments of the previous sections are made very small the phenomena are entirely altered. The images formed by a pin-hole camera are sharper the smaller the hole is, but also feebler. The diameter of the hole, however, cannot be reduced below about 1/100 mm., otherwise the images become less sharp again. The explanation of the formation of the images given on p. 8 therefore fails when the opening is less than 1/100 mm. in diameter. Further, if the screen which serves as an obstacle to the rays of light is made very small, a bright patch may actually occur at the place where we should expect to find the umbra, i.e. complete darkness. The case of a small opening has already been studied in Vol. II, p. 244 (fig. 28). The reason for these phenomena was brought out clearly there; they are due to **diffraction**, the capacity of waves for passing round an obstacle and reinforcing one another in the rear of the obstacle by superposition (Vol. II, p. 241; present volume, p. 14).

We shall now discuss these diffraction phenomena, which form direct evidence in favour of the wave theory of light, in more detail

If light does consist of waves, it should exhibit phenomena similar to those shown by the water waves in fig. 9 (Vol. II, p. 215), where straight wave fronts are incident from the right on a screen with a gap in it. The waves are not only propagated through the opening in their original direction, but they spread out behind the opening as though the opening were itself the centre of a system of concentric waves. The amplitude of the concentric waves is greater in the direct line of advance behind the opening than at either side; the smaller the opening in the screen, the smaller is this difference, and the larger the opening, the more marked does the difference become

In actual fact the following experiments * show that light does exhibit phenomena similar to those which may be observed directly with water waves.

(1) We produce a light as like a point source as possible by illuminating a small opening in an opaque screen from the rear by sunlight or an arc lamp, or by using the light from the crater of the arc lamp directly. At a distance of about 1 metre we set up a screen which has had a hole pricked in it by a sharp

* The fundamental experiments on diffraction were first carried out by the Jesuit FRANCESCO GRIMALDI (1618–1663). He held firmly to the corpuscular theory of light and could find no explanation for the phenomena, although he knew that if light is added to the light received by a luminous body, the body may become darkened. On becoming acquainted with GRIMALDI's discovery, HOOKE (1665) proposed to explain it by the assumption that light is a wave motion.

needle, and we receive the light passing through this hole on a white screen placed a metre farther on. We observe that the resulting spot of light is considerably larger than it would be according to the geometrical diagram and that its edges are blurred. This phenomenon is very reminiscent of the water waves passing through an opening in a screen.

(2) If we look at a point source of light through a small needle-hole in a piece of paper, the source of light looks larger than when it is observed directly, and, moreover, appears to be surrounded by several more or less coloured circular rings.

(3) If we look at a point source of light through a narrow slit, the source of light appears greatly broadened in the direction at right angles to the slit, but not in the direction of the slit. In the broadened part we observe several bright and dark bands or fringes which run parallel to the slit and are more or less coloured. If we use a tapering slit, these bands are close together at the broad end of the slit and far apart at the narrow end. Any irregularities in the edges of the slit are betrayed in an exaggerated form by irregularities in the diffraction pattern. These bands arise from the superposition of wave motions (p. 15).

Fig. 15 (Plate I) is a natural-size photograph of the shadow of a tapering slit with straight sides, taken by monochromatic light (wave-length $\lambda = 0{\cdot}46\ \mu$). The slit was at a distance of 24·17 metres from a source of light about 1 mm. in diameter and the photographic plate was at a distance of 15·47 metres from the slit. The geometrical shadow of the slit is indicated by dotted lines in the photograph. We see that the edge of the shadow is displaced a long way from the geometrical shadow and that interference bands occur, the latter being farther apart the smaller the breadth of the slit (Arkadiew). Similar experiments have been made with X-rays by B. Walter (1924).

(4) If we look directly at a source of light but hold a wire about 0·2 mm. in diameter just in front of the eyes, or if we hold the wire at arm's length and look at it through a strong hand-lens, we see a number of bright and dark bands in the blurred shadow of the wire, running parallel to the wire. The central portion of the shadow, i.e. that part which according to the geometrical diagram should be darkest, is bright, and the black bands occur at equal distances on either side. The last phenomenon may also be displayed objectively by receiving the shadow of the thin wire on a white screen at a distance of about 2 metres. Here again the very middle of the shadow is bright.

Fig. 16 (Plate I) is the photograph of the shadow of a needle 1·97 mm. in diameter, taken by monochromatic light (wave-length $\lambda = 0{\cdot}46\ \mu$). The needle was illuminated by a slit $0{\cdot}7 \times 4$ mm. parallel to the needle; the needle was 24·17 metres from the slit and the photographic plate 15·47 metres from the needle (Arkadiew).

Similar phenomena are observed if we look at a bright and distant source of light with the eyes nearly closed, so that the eyelashes form a "grating" in front of them.

(5) If we place a small opaque screen with a sharp edge in the path of the rays from a distant point source, this screen does not throw a sharp shadow on another screen, but the bright region fades gradually into the dark region, and in front of the region of geometrical shadow the illuminated portion of the shadow contains a number of more or less coloured bands parallel to the edge of the geometrical shadow (fig. 16).

Theory of the Diffraction Phenomena. In Vol. II (p. 241) it has already been indicated how these diffraction phenomena can be explained quantitatively on the basis of Huygens' principle and its extension by Fresnel. The case of a *small circular opening*, i.e. the

above-mentioned case of the very small opening of a pin-hole camera, was discussed there.

We now go on to discuss the case of a small circular obstacle.

Small Circular Obstacle.—In fig. 17 B denotes a circular obstacle with its edges at NN illuminated by parallel rays of light LL. According to the geometrical theory there would be a cylindrical region of shadow behind the obstacle, with its axis OP along the normal to the obstacle at its centre. In actual fact, however, the middle of this geometrical shadow is always bright when the obstacle NN is small in comparison with the distance a. In the figure the dimensions have necessarily been drawn incorrectly, as the distance a is actually several metres but the diameter of the obstacle only a few wave-lengths.

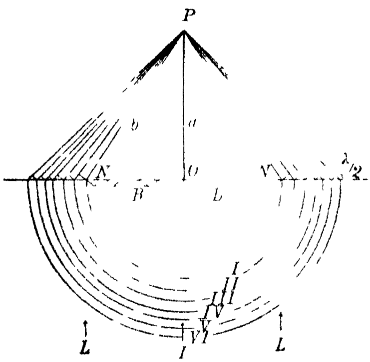


Fig 17.—To illustrate the theory of diffraction at a small obstacle

If we imagine the plane of the obstacle extended, this extension forms a wave surface for the parallel rays of light LL. Just as was described in Vol. II (p. 241), we now construct spherical surfaces about a point P lying on the axis, the innermost of these spheres passing through NN, the rim of the obstacle. If the radius of this sphere is b, the following spheres must have radii $b + \lambda/2$, $b + 2\lambda/2$, $b + 3\lambda/2$, &c. These spherical surfaces cut elementary Fresnel zones out of the wave surface; their boundaries (circles) are half shown in the figure, and they are to be imagined as rotated through 90° in the plane of the obstacle. By Vol. II, p. 242, the elementary zones all have the same area and emit waves in such a way that the waves from two neighbouring zones differ by exactly half a vibration at the instant when they reach the point P. The same considerations apply to these elementary zones as to an infinite plane wave (Vol. II, p. 243). If we denote the effect of zone I at P by z_1, that of zone II by z_2, and so on, the total effect of all the zones at the point P is

$$z = z_1 - z_2 + z_3 - z_4 + \ldots \textit{ad infinitum}$$

We write this series as follows

$$z = \tfrac{1}{2}z_1 + (\tfrac{1}{2}z_1 - z_2 + \tfrac{1}{2}z_3) + (\tfrac{1}{2}z_3 - z_4 + \tfrac{1}{2}z_5) + \ldots$$

As was explained on p. 243 of Vol. II, each bracket becomes equal to zero, so that the resultant effect reduces to

$$z = \tfrac{1}{2}z_1.$$

Hence:

If a pencil of parallel rays falls normally on a small circular obstacle, the illumination at a point behind the obstacle on its axis is the same as if half the elementary Fresnel zone nearest the rim of the obstacle were alone effective.

The light accordingly spreads out even behind the obstacle. This, of course, happens no matter what the size of the obstacle. If the diameter of the obstacle NN is $2r$, the area of the first elementary zone, by considerations similar to those

of Vol. II, p. 242, is $\pi\lambda b$. The intensity at P will be proportional to this area, and will also decrease proportionally to $b^2 = NP^2$ according to the law of inverse squares. According to FRESNEL, the elementary disturbances which reach P from the first elementary zone also decrease as the angle PNO decreases. As r increases, the light reaching the point P from the first elementary zone therefore decreases more rapidly than the expression $\pi\lambda b/b^2$ or $\pi\lambda/\sqrt{(a^2+r^2)}$. If the obstacle is relatively small (r small compared with a), the angle PNO (fig. 17) may be regarded as a right angle; the light reaching P from the first elementary zone is greater the smaller a is. If, however, the obstacle is relatively large (r is not negligible compared with a), P receives less light the greater r (and also a) is compared with λ.

At a sufficient distance behind a small obstacle, therefore, there is no longer a proper shadow (when parallel light is used); at any one distance, however, the shadow is more clearly developed the larger the obstacle is.

If we imagine the point P laterally displaced, e.g. to the right, the figure becomes a great deal more complicated; we can see at once, however, that the brightness must be increased, as then the elementary zones adjacent to the right-hand edge of the obstacle become broadened much more than the zones on the left-hand side are narrowed.

Photographs of the shadows of two small metal obstacles are shown in fig. 18 (Plate I). The diameters of the obstacles were such that the first covered *one* elementary Fresnel zone, the second *two* elementary Fresnel zones. In both figures the white spot in the centre is noteworthy as showing that diffracted light is present at the very centre of the geometrical shadow.* The photographs of fig. 18 and of fig. 28 (Vol. II, p. 245) were taken with monochromatic light of wave-length $\lambda = 0{\cdot}46\ \mu$. The distance of the point source of light from the obstacle or opening was 27·77 metres, and the distance of the photographic plate from the obstacle or opening 11·7 metres (ARKADIEW).

Narrow Rectilinear Obstacle (Thin Wire).—We may regard NN in fig. 17 as the cross-section of a narrow rectilinear obstacle extending to an infinite distance at right angles to the plane of the paper. The elementary Fresnel zones I, II, &c., then become *narrow rectilinear strips parallel to the obstacle* B. These elementary zones, however, no longer have equal areas, the area slowly decreasing as the number of phase jumps increases.

This decrease in the radiating surface leads to a decrease in the disturbances sent by the individual zones to a point in the geometrical shadow behind the centre of the obstacle, just like the decrease in amplitude of the individual disturbances due to the elementary zones which we discussed in Vol. II (p. 243). If, as on that occasion, z_0, z_1, z_2 are the disturbances produced by the elementary zones at the point in question, the z's slowly decrease as the number of the

* This seemingly absurd result was deduced by POISSON (Vol. II, p. 93) from Fresnel's diffraction theory; he found that the light must penetrate into the middle of the geometrical shadow of a round obstacle just as if there were no obstacle in the path of the rays. POISSON, who thought himself obliged to defend the then current Newtonian idea of the corpuscular nature of light against the wave theory, regarded his deduction as an objection to FRESNEL, who was a sponsor for the wave theory, and thus contributed to the victory of the latter on the phenomenon being actually observed.

elementary zone increases, on account both of the increase in distance and of the decrease in area. As before, we may assume that the disturbance z_n is the

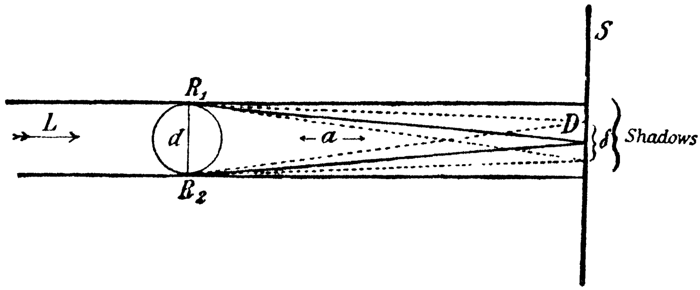


FIG. 19.—To illustrate the theory of diffraction at a wire (measurement of wave-length)

arithmetic mean of the disturbances due to the neighbouring zones. Every point behind the centre of the obstacle is therefore illuminated as if only the first half of the two elementary strips I lying to the right and to the left acted on it; the effects of the other elementary strips accordingly cancel one another, just as we deduced in the case of the small circular obstacle (fig. 17).

Measurement of Wave-length by Diffraction at a Wire.—This case is of special practical importance inasmuch as it enables us to determine the wave-length of the light used fairly accurately by very simple means. For the two edges of the narrow obstacle or the two lines of contact of a pencil of parallel light rays with a cylindrical wire act just as if these edges were coherent sources of light; each side has the same effect at a point in the shadow of the obstacle or wire as the half of the first rectilinear Fresnel zone immediately adjacent to the edge.

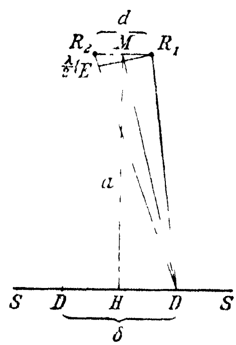


FIG. 20.—To illustrate the theory of Fresnel's mirror experiment (see also p. 174)

Let the small circle in fig. 19 represent the cross-section of the wire; its diameter d can be measured by means of a screw gauge. The wire is illuminated from the left by parallel rays travelling in the direction L from a source of light (slit). From the edges R_1, R_2 of the wire two systems of light waves travel into the region of the geometrical shadow; these systems are coherent, as they start from the same source of light and agree in wave-length and phase. It follows from this that the centre of the shadow on the screen S must be bright (fig. 21), for the centre is equidistant from the two edges, so that here the light waves meet in the same phase. At the point D there is a dark band, provided the difference of its distances from the two edges of the wire is equal to half a wave-length or an odd multiple of half a wave-length. If we call the distance between two adjacent dark bands δ, i.e. the distance of the central bright band from the first dark band $\delta/2$, the distance of the screen from the wire a, the thickness of the wire d, and the wave-length of the light λ, it follows that if we describe a circle

with centre D and radius DR_1 cutting DR_2 at E (fig. 20), we have $R_2E = \lambda/2$. Owing to the great distance of the screen SS from R_1 and R_2, we may regard R_1E as a straight line perpendicular to R_2E and hence R_1R_2E as a right-angled triangle. If we join MD, it follows from the equality of their angles that the triangles R_1R_2E, MDH are similar. Hence $R_1R_2 : R_2E = MD : DH$, and as (owing to the great distance of the screen from the source of light as compared with the distance between the bands) we may put $MD = MH$, we have

$$\lambda = \frac{d\delta}{a}.$$

Fig. 21 shows the diffraction pattern projected on a screen carrying a scale.

Of the bands which are formed only the central one is pure white, the bands further removed from the centre exhibiting coloured edges. To investigate this phenomenon we observe the bands simultaneously through a glass coloured red in its upper half and blue in its lower half (cf. fig. 11, p. 183). We find that the two systems of bands now consist of red and black or blue and black bands, and that the red bands are farther apart than the blue ones. Hence, according to the above formula, the wave-length of red light is greater than that of blue light (Vol. III, p. 646). At a greater distance from the centre the superposition of the systems of bands for the various colours gives rise to white bands with coloured edges (Chap. VII, § 2, p. 154).

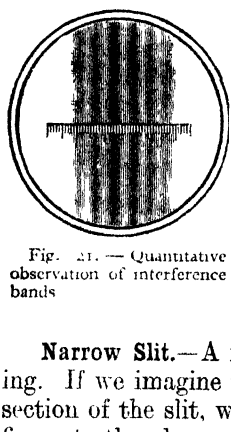


Fig. 21. — Quantitative observation of interference bands

Narrow Slit.—A narrow slit behaves essentially like a small opening. If we imagine that fig. 27 (p. 244 of Vol. II) is a vertical cross-section of the slit, we can apply most of the discussions based on this figure to the phenomena exhibited by a slit.

The elementary Fresnel zones I, II, III, &c., however, are now rectangles (loc. cit.) with their long sides parallel to the slit and their short sides given by OA_1, A_1B_1, B_1C_1, &c. The areas of these elementary zones are no longer equal but decrease as the number of the zone increases. The disturbance at a point P behind the centre of the slit, which, to a first approximation, is proportional to the area of the elementary zones, therefore decreases as the number of elementary zones emitting light increases. The disturbance arising from the first elementary zone preponderates over all the rest. In any case, therefore, the central region behind the slit is bright. But according as the first elementary zone alone, or the two first zones, the three first, &c., contribute to the disturbance—these cases may occur as the distance a of the point P from the slit diminishes—the brightness observed along the axis P_1O goes through a maximum for one elementary zone, a minimum for two, a second maximum for three, and so on. Thus there is an alternating sequence of dark and bright bands on either side. The case when P is at an infinite distance is discussed in detail in Chapter VIII, § 8 (p. 198).

Diffraction at a Single Edge.—In fig. 22 R represents an obstacle unbounded on the left, with the rectilinear edge N, on which the

pencil of *parallel* rays of light LL is incident. The white screen S is set up behind this opaque obstacle. NP_0 is the boundary of the geometrical shadow.

According to Huygens' principle all the surface elements of the plane wave surface W which just grazes the diffracting edge N contribute to the disturbance at any point P at a distance a from the wave surface. Owing to the variation of the distances from the points of the wave surface to P, however, these disturbances arrive at P with varying phases. From P we draw the rays $PQ_0 = a$, $PQ_1 = a + \lambda/2$, $PQ_2 = a + 2\lambda/2, \ldots, PQ_n = a + n\lambda/2$ to the points $Q_0, Q_1, Q_2, \ldots, Q_n$ of the wave surface. Then

$$Q_0Q_n^2 = PQ_n^2 - PQ_0^2 = (a + n\lambda/2)^2 - a^2 = na\lambda,$$

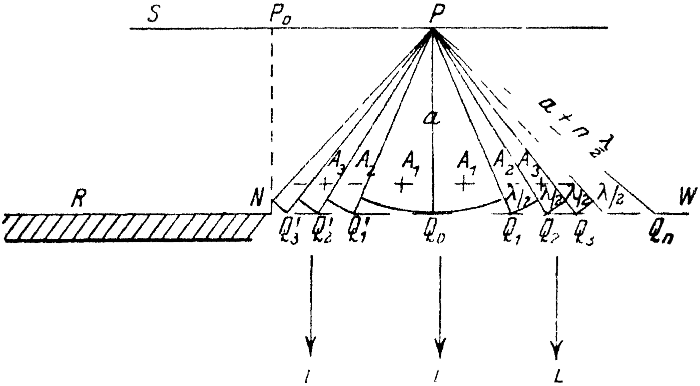


Fig. 22. - Diffraction at a straight edge

provided we may neglect $n^2\lambda^2/4$. The disturbances produced at P by the strips of surface of breadth Q_0Q_1, Q_1Q_2, Q_2Q_3, . . . parallel to the straight edge N must be approximately proportional to these breadths. Corresponding to every ray starting from the strip of breadth Q_0Q_1, there can be found one of opposite phase in the neighbouring strips. The disturbances which arrive at P from Q_0Q_1, Q_1Q_2, Q_2Q_3, &c., therefore have alternating signs. Let the disturbance due to the strip of breadth Q_0Q_1 be A_1, that due to Q_1Q_2, A_2, and so on. Then Q_0Q_n gives rise to the disturbance $S_0 = A_1 - A_2 + A_3 - A_4 + \ldots + A_n$. Here A_n is proportional to $Q_{n-1}Q_n = Q_0Q_n - Q_0Q_{n-1} = \sqrt{(na\lambda)} - \sqrt{\{(n-1)a\lambda\}}$, i.e. as n increases the amounts A_n diminish to zero. The series

$$S_0 = A_1 - A_2 + A_3 - A_4 \ldots,$$

being a series of terms alternately positive and negative, therefore converges towards a finite limit.

The disturbance at P only has the value S_0 when P is at P_0, the geometrical shadow of the edge N. If we now let P move from P_0 towards the right through the radiation field, the disturbance at P at first increases, for P is then also illuminated by points lying to the left of Q_0. The points of the strip $Q_0Q_1' = Q_0Q_1$ give rise to disturbances in the same phase as those due to points in the strip Q_0Q_1. Hence if

$$P_0P = Q_1'Q_0 = Q_0Q_1 = \sqrt{(a\lambda)},$$

the point P has reached a point of maximum disturbance $S_1 = S_0 + A_1$. As P advances farther it must reach points of lesser disturbance, as to the left of Q_1' there lie points which give light with the same phase as that given by the points of Q_1Q_2, i.e. light whose phase is opposite to that of A_1, the greatest contribution to the disturbance. The light intensity at the points traversed by P must accordingly diminish until P_0P attains the length of $Q_0Q_2 \approx \sqrt{(2a\lambda)}$. At the distance $P_0P \approx \sqrt{(2a\lambda)}$ there is minimum disturbance of amount $S_0 + A_1 - A_2$. If we let P move on farther, the same considerations for any value x of P_0P lead to a succession of minima and maxima. For $x = 0$ we have the disturbance S_0, for $x = \sqrt{(a\lambda)}$ the disturbance $S_0 + A_1$, a maximum, for $x = \sqrt{(2a\lambda)}$ the disturbance $S_0 + A_1 - A_2$, a minimum, for $x = \sqrt{(3a\lambda)}$ the disturbance $S_0 + A_1 - A_2 + A_3$, a maximum, and so on. Summarizing, we have the following:

The region outside the geometrical shadow exhibits dark and bright interference bands which are parallel to the edge of the obstacle and whose intensities become more and more equal as the distance from the geometrical shadow increases.

The bright bands are at distances $x_n = \sqrt{(na\lambda)}$, where $n = 1, 3, 5, \ldots$, from the edge of the geometrical shadow, and the dark bands at distances $x_n = \sqrt{(na\lambda)}$, where $n = 2, 4, 6, \ldots$

Using more powerful methods of mathematical investigation, Fresnel deduced an approximation formula,

$$x_n = \sqrt{\{a\lambda(n - \tfrac{1}{4})\}},$$

where $n = 2, 4, 6, \ldots$; this is somewhat more accurate, particularly for the position of the first dark band.

Within the geometrical shadow there lie corresponding maxima and minima; thus to the left of P_0 at a distance $x = -\sqrt{(a\lambda)}$ the disturbance is $S_0 - A_1$. As, however, A_1 determines the sign of S_0, being the term of greatest absolute value in the series for S_0, the disturbance at this distance must have the opposite sign, i.e. the opposite phase to that at P_0. Hence between $x = 0$ and $x = -\sqrt{(a\lambda)}$ there is a minimum where the absolute value of the disturbance is zero, but at the distance $x = -\sqrt{(a\lambda)}$ there is a maximum, at which the light has the phase exactly opposite to that at the point $x = +\sqrt{(a\lambda)}$. Maxima with their phases successively alternating follow at the distances $x_n = -\sqrt{(na\lambda)}$, where $n = 1, 2, 3, 4, 5, \ldots$ The distances between the interference bands are therefore less than they are outside the geometrical shadow, and the difference of intensity between the light and dark bands diminishes much more rapidly.

Divergent Light.—If the incident light is not parallel but comes from a narrow slit S (fig. 23) parallel to N, the edge of the obstacle, and at the distance $SQ_0 = r$ from the plane of the obstacle, considerations similar to those above apply. We have only to replace a by $ra/(r + a)$ in the final formula for the position of the interference bands.

Proof.—The elementary zones now lie on the curved surface of a cylinder; $PQ_n' = a + n\lambda/2$. Then

$$Q_n'P^2 = \left(a + \frac{n\lambda}{2}\right)^2 = Q_n'L_n^2 + (a + Q_0L_n)^2,$$

or

$$an\lambda = Q_n'L_n^2 + 2aQ_0L_n + Q_0L_n^2 \quad . \quad . \quad . \quad (a)$$

if we may again neglect $(n\lambda/2)^2$ compared with $an\lambda$. Further,

$$Q_nS^2 = r^2 = Q_n'L_n^2 + (r - Q_0L_n)^2$$

or

$$0 = Q_n'L_n^2 - 2rQ_0L_n + Q_0L_n^2.$$

Subtraction from the previous equation gives

$$an\lambda = 2Q_0L_n(a + r) \text{ or } Q_0L_n = \frac{an\lambda}{2(a + r)}.$$

Hence, again neglecting certain terms, we have, by equation (a),

$$Q_n'L_n^2 = an\lambda - 2a \cdot \frac{an\lambda}{2(a+r)} = \frac{arn\lambda}{a+r}$$

and
$$PL_n = a + \frac{an\lambda}{2(a+r)} \approx a.$$

Since $x_n : Q_n'L_n = a : PL_n$, we then have

$$x_n = Q_n'L_n = \sqrt{\left(\frac{arn\lambda}{a+r}\right)}.$$

As $r \to \infty$ this tends to the formula $x_n = \sqrt{(an\lambda)}$ above.

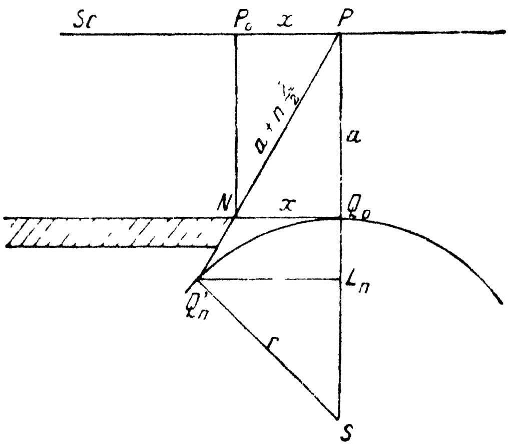


Fig. 23.—Diffraction at a straight edge with divergent light

Experiment.—In fig. 24 (Plate I) photographs of the shadows of the edges of various bodies in divergent light are shown. The uppermost of the four photographs is the shadow of the edge of a sharp razor-blade, the second the shadow of the edge of a glass rod 7·8 mm. across, the two others being shadows of bodies with a radius of curvature of 40 metres. In all four cases the diffraction bands appear at the same intervals. The photograph was taken with light of wave-length $\lambda = 0{\cdot}46\ \mu$ coming from a slit parallel to the diffracting edge. The distance r between the source of light and the diffracting body was 24·17 metres and the distance a between the diffracting body and the photographic plate was 15·47 metres. Below the photographs is a graph of the theoretical distribution of intensity (Arkadiew).

White Light.—In our discussion above we have always taken for granted that the light used was monochromatic. Now the breadth of the diffraction bands depends on the wave-length. In particular we found in our last proof (see above) that the distance of the bands from the edge of the geometrical shadow is $x_n = \sqrt{(na\lambda)}$. Hence if we use a mixture of lights, e.g. white light, every colour gives rise to its own diffraction bands, which accordingly overlap. Hence diffraction

bands, as well as other interference bands, are coloured when white light is used, each bright band being edged with blue on the inner side and with red on the outer side. At greater distances, where the bands follow in rapid succession, the separate components become mixed and form white.

Effect of the Material of the Obstacle.—Hitherto we have simply regarded the obstacle as opaque. Light waves, however, consist of electromagnetic vibrations. Hence it is not immaterial what electric or magnetic conditions exist at the edge of the obstacle or opening, whether it is a metal or an insulator, &c. These effects may be compared say to the effect on water waves of an obstacle made of soft or elastically yielding material. Here we merely mention these problems which are very difficult to attack theoretically

The phenomena adduced above may suffice to acquaint the reader with the essential properties of light as a wave motion. These will be treated in greater detail below once we have studied more refined instruments for the investigation of the phenomena (Chap. VIII p. 174).

3. The Distribution of the Light from a Source.

The distribution of the light emitted by ordinary light sources differs in different directions. Hence if a source of light is to be completely specified, its intensity must be investigated separately for a variety of directions. Many sources of light are bodies of revolution, so that we can content ourselves with an investigation of the distribution of light in a single meridian.

An electric bulb can be placed in any position and then be rotated so that all its points successively send out light horizontally, i.e. parallel to the photometer bench. All other sources of light, and in particular those involving flames, can only be used in one definite position, so that special apparatus has to be employed to turn the light emitted by the source from the directions under investigation into the horizontal direction; this can be done by arrangements of mirrors which can be adjusted in various directions.

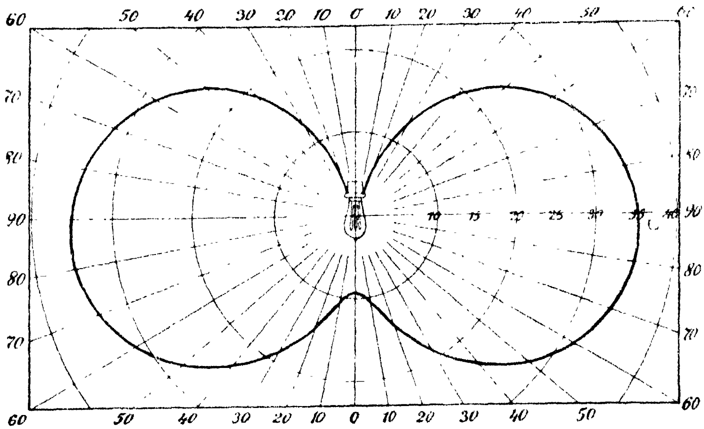


Fig. 16.—Distribution of light from an Osram lamp

To obtain a rapid idea of the distribution of light, the intensities obtained for the different directions are plotted as the radii vectors of a system of polar co-ordinates and the ends joined by a curve. The distribution of the light from an electric bulb is shown in fig. 16.

A curve of this kind may be used to calculate the whole luminous flux from a source of light. If the source L (fig. 17) is at the centre of a unit sphere, the latter may be divided up by meridians and circles into elements of surface $d\omega$, which are numerically equal to the corresponding space angles $d\omega$. By § 1, p. 21, the luminous flux through this element of surface is given by the product of the magnitude of the area and the intensity I. If we form this product $I\,d\omega$ for all the elements of the space angle, we obtain the total luminous flux of the source of light by summation: that is, $F = \int I\,d\omega$. Dividing this expression by 4π, we obtain the **mean spherical candle-power** or **average candle-power** of the source of light. The luminous flux may also be measured for the upper or lower hemisphere alone; on division by 2π this gives the **mean hemispherical candle-power** (upper or lower).

The calculation of the mean intensity in the way indicated above is a troublesome and lengthy task; attempts have therefore been made to determine the whole luminous flux by means of a *single* measurement. This is made possible by the use of the **Ulbricht globe** or **sphere photometer.** This consists of a large hollow sphere from 1 to 3 metres in diameter (fig. 18), whose inside surface is matt and painted pure white. If the source of light under investigation is placed in a sphere of this kind, the light which it emits in all directions is reflected diffusely many times, and every element of the inner surface of the sphere is equally illuminated. Then we have merely to measure the illumination of an element of surface, which is protected from direct illumination by the screen S, through a hole for observation O in order to derive a correct result for the total luminous flux. The Ulbricht globe so to speak integrates the individual luminous fluxes and is therefore referred to as an **integrating photometer.**

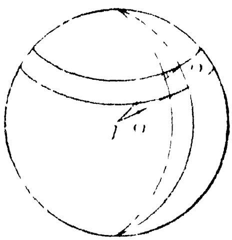


Fig. 17.—Total luminous flux

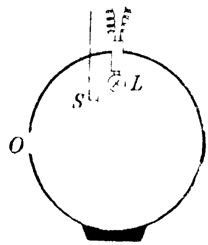


Fig. 18.—The Ulbricht globe

Finally, the question of the illumination produced by a source on a given surface is of great importance. The illumination can either be measured directly by an illumination photometer (p. 27) or calculated from the distribution of the light of the source using Lambert's laws, the height and distance of the source and the angle at which the light rays fall on the surface being known.

4. The Mechanical Equivalent of Light.

Light consists of electromagnetic energy propagated with the speed of light. When this falls on the retina of the eye (p. 113) it is transformed into the sensation of light. The measure of light sensation is the luminous flux passing into the eye, measured in lumens. Or the power of the source may be measured in watts or calories per second, e.g. by letting the light fall on a suitably constructed calorimeter (or on a black body (Vol. III, p. 247)), and measuring the rise in temperature. We can thus find how many watts or calories per second correspond to one lumen. Here, however, it is to be noted that the eye is not equally sensitive to all wave-lengths. The continuous curve in fig. 19 is the curve of sensitiveness of the eye for bright light; it has a maximum for the wave-length 555 $m\mu$ (yellowish-green). This light is also the strongest constituent of sunlight as received at the earth's surface, a fact which has led to interesting deductions as to the course of evolution.

The number of watts which in the form of light of the wave-length to which the eye is most sensitive represents a luminous flux of 1 lumen.

is called the **mechanical equivalent of light** (M). Experiments give

$$M = 0{\cdot}00133 \frac{\text{watt}}{\text{lumen}} \pm 5 \text{ per cent.}$$

For light of a different wave-length we have also to take into account the curve of sensitiveness of the eye, which is given below.

Note.—For faint light (twilight) the eye has a different curve of sensitiveness, which is shown dotted in fig. 19 (the *Purkinje effect*). Under certain circumstances this must also be taken into account in the calculation.

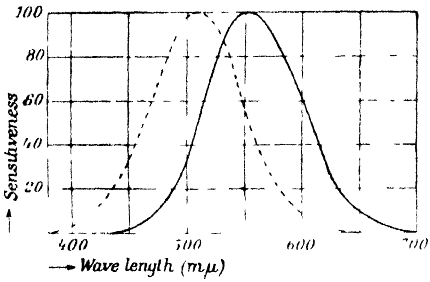


Fig. 19.—Curve of sensitiveness of the eye for bright light (continuous) and for feeble light (dotted).

Most of the sources of light used at present depend on the production of light through heat (temperature radiation *), and in addition to visible light emit very large amounts of invisible radiation, mostly on the long-wave side. Of the visible light, again, only a fraction is utilized by the eye, in accordance with its curve of sensitiveness. Fig. 20 exhibits the relationships for a body heated to 3000° abs., corresponding more or less to a modern electric bulb. The energy of the radiation of the body is plotted as ordinate against the wave-length as abscissa. The shaded area denotes the portion of energy which is radiated as visible light, the black area the fraction corresponding to the sensitiveness of the eye. We see how uneconomical this method of light production is. Pirani has recently succeeded by an entirely different method (arc discharge in sodium vapour) in obtaining extremely high yields of light, nearly reaching the mechanical equivalent of light.

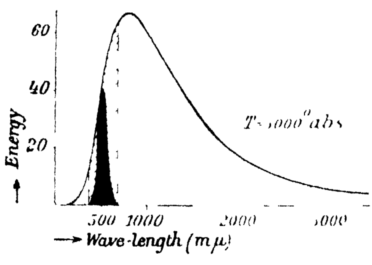


Fig. 20.—Energy of total radiation (area of continuous curve) compared with the energy of visible radiation (shaded) and the brightness perceived by the eye (black). Radiation of a black body at 3000° abs.

A measure for the yield of light is given by the fraction of energy supplied which is transformed into visible light, i.e. the intensity in international candles produced by one watt. The yield is accordingly measured in candles/watt. The reciprocal watts/candle gives the power which must be supplied to produce an intensity of one candle-power.† Such values have already been given in Vol. III (pp. 259, 339, and 356); see also *Temperature radiation*, Vol. V.

* See Vol. V.

† [In practice the "efficiency" of lamps is expressed in lumens per watt. Modern tungsten filament lamps have efficiencies ranging from 8 lumens per watt for the small sizes up to 30 lumens per watt for the higher wattage types.]

CHAPTER III

Geometrical Optics: Reflection and Refraction

1. Principles of Geometrical Optics.

Often it is unnecessary to take account of the wave nature of light (Chap. I, §§ 3, 5, pp. 6, 9). Thus, for example, in discussing the images formed by the pin-hole camera, we may assume that the light rays still remain straight lines when passing through the opening. As was shown in § 6 of Chapter I (p. 11), this assumption is justified only when the opening is uniformly occupied by the light and is large compared with the wave-length of the light. These conditions, however, are satisfied in a very large number of cases occurring in practice. In these cases, therefore, the rays of light may be treated like geometrical straight lines, and many problems may thereby be readily solved, some of which are very difficult to discuss on the basis of the wave theory. This method of treating optical problems is known as **geometrical optics** (or ray optics), while the method where the wave properties of light are taken into account is known as **physical optics** (or wave optics). The following four chapters are based on the above assumptions. We shall begin by considering the radiation emitted from a very small element of surface. In practice the crater of an arc lamp or a small opening in an opaque screen used to cut down the radiation from a source of light may be regarded as fulfilling this object.

In geometrical optics the course of the rays is independent of direction, that is, any particular path followed by the rays may also be traversed in the reverse direction (*law of the reversibility of the path of the light*). Thus, e.g., in reflection the incident ray and the reflected ray are interchangeable. See also pp. 37, 51.

2. The Plane Mirror.

Applications.—The action of the mirror was described on p. 7. We shall now mention some applications.

The ordinary plane mirror is used either when we want to make a given light ray travel in a prescribed direction or when we wish to calculate the amount through which a mirror has rotated from the change in direction of a light ray.

The **heliostat** or **coelostat** * is used to bring rays of light from the sun or stars into the observation-room through an opening in the wall. According as the heliostat is turned by hand or by clockwork so as to follow the apparent motion of the sun, it is called a *hand heliostat* or *clockwork heliostat*. In clockwork heliostats the axis of rotation of the mirror is set parallel to the earth's axis; the mirror rotates once in twenty-four hours, and always reflects the sun's rays in a direction parallel to the earth's axis. The rays are then sent in any desired direction by means of a second mirror. The **heliograph** (fig. 1, Plate II) used for signalling works in exactly the same way. The mirror on the left, driven by clockwork, sends sunlight in a prescribed direction to the actual signalling mirror on the right; signalling can then be continued for long periods without the necessity for adjusting the apparatus as the position of the sun in the sky varies.

The **reflecting goniometer** (shown in outline in fig. 2 and in perspective in fig. 3, Plate III) is used to find the angle between two surfaces (e.g. of a crystal or prism). The goniometer consists of a divided circle K, at the centre of which a small table T is fixed so as to be free to revolve. To the table there is rigidly fixed a rotating arm (not shown in fig. 2) carrying a zero mark and a vernier from which the rotation of the table may be read off. A tube S provided with a slit and a convex lens (the **collimator** †) is also fixed to an arm rotating about an axis coinciding with the axis of the table and passing through the centre of the divided circle. The amount through which this arm rotates can also be read off on the divided circle. The position of the collimator slit relative to the lens is such that the rays of light entering the slit leave the collimator lens as a parallel pencil of rays (p. 63). Finally, an arm carrying a telescope F is also fixed to the axis of the goniometer so as to rotate about the latter; its rotation can also be read off on the divided circle. The telescope is so adjusted that on looking in a straight line through the telescope and the collimator, one sees a clear image of the slit. The telescope usually contains cross-wires (p. 137), so that the direction of the telescope can be accurately established.

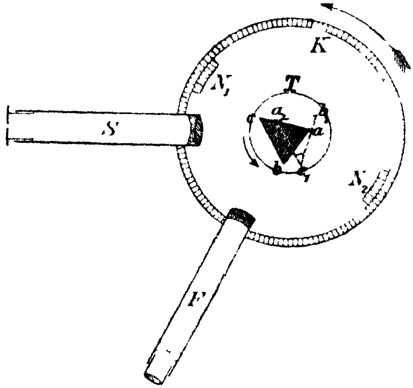


Fig. 2.—Plan of a reflecting goniometer

In order to measure the angle of a prism with the reflecting goniometer, we set the prism *abc* on the table T (fig. 2) so that the edge *c* of the prism is parallel to the axis of rotation of the goniometer. A pencil of rays passing through the collimator S falls on the surface *bc* of the prism and is reflected there. The telescope T is then adjusted so that the reflected ray of light falls exactly on the

* Gr., *hēlios*, the sun; Lat., *status*, fixed; Lat., *coelum*, the sky.

† Lat., *limare*, to polish, to investigate carefully, from *lima*, a file. The word *collimare* was erroneously formed, in place of the correct term *collineare*, to bring into a straight line, owing to a wrong reading in CICERO; the tube should really be called a *collineator*. The term collimator was originally applied to the auxiliary telescope attached to large telescopes, its use being subsequently extended to the apparatus described above.

cross-wires. The table and prism are next rotated so that the light ray entering the collimator is reflected into the telescope by the second face of the prism, *ac*. The new position of the prism is shown by dotted lines. The angle through which the table has been rotated is read off on the divided circle: it is equal to the supplement of the angle of the prism to be determined.

The **mirror and scale** (fig. 4) is used to measure the rotation of a small mirror accurately (e.g. of a mirror which is fixed to the suspended magnet of a galvanometer and rotates with it). At a measured distance *a* the telescope F is set up so that the normal to the mirror coincides with the axis of the telescope. We then observe the zero mark O of a scale MM set up at right angles to the telescope and usually just above or below the telescope objective, by reflection at the mirror.

If the mirror S rotates through the angle φ, the normal to the mirror rotates with it, and we then see a point A of the scale such that $\angle OSA = 2\varphi$. The

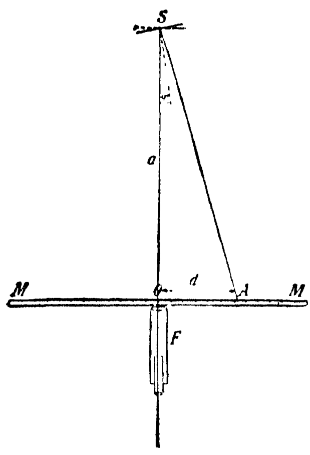


Fig. 4.—Mirror and scale

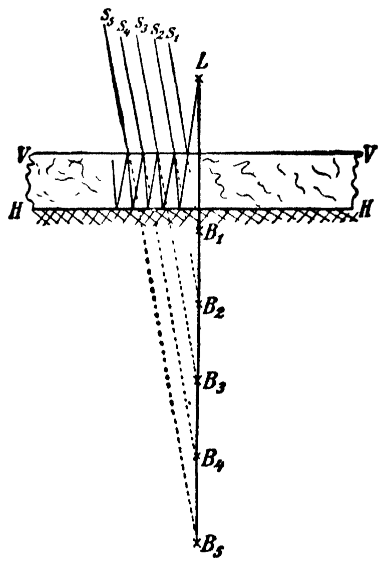


Fig. 5.—The succession of images formed by a plane mirror

position of the observed point A is determined by the distance $OA = d$. We can observe $\tan 2\varphi = d/a$ very accurately, as both *d* and *a* can be measured accurately. If the angle φ is small the tangent of the angle may be replaced by the angle itself (in radians); for small deflections, therefore, the observed linear deflection is proportional to the angle of deflection.

Combinations of Mirrors. Parallel Mirrors.—A large number of the mirrors in use are of glass silvered at the back. When a ray of light falls on a mirror of this type, part of it is reflected at the front surface of the glass and part of it penetrates into the glass. The latter part is almost entirely reflected at the silvered face, i.e. it returns to the front surface and is there again divided into two parts, one returning through

the glass and the other passing out of the glass. The part remaining in the glass is again reflected at the back surface and again subdivided at the front surface. The reflected portion of the ray traverses the layer of glass in both directions for the third time, and the process is repeated until the residue of light becomes so feeble that it can no longer be detected.

The process is illustrated in fig. 5. The splitting-up of the rays is shown for a ray incident almost normally on the front surface (VV) of the mirror. HH is the silvered rear surface of the mirror. The parts of the rays which emerge from the front surface are denoted by S_1, S_2, &c. (Meanwhile we do not stop to consider the change in direction of the ray on entering the glass.) We see that the repeated reflection must give rise to a large number of images of the source of light L; these lie one behind the other, so that they are not visible separately if we look at the mirror in a direction exactly at right angles to it. (Owing to the refraction in glass which we mentioned above, the distance between the successive images B_1, B_2, &c., is equal to three-quarters of the thickness of the glass: the brightest image is B_2.)

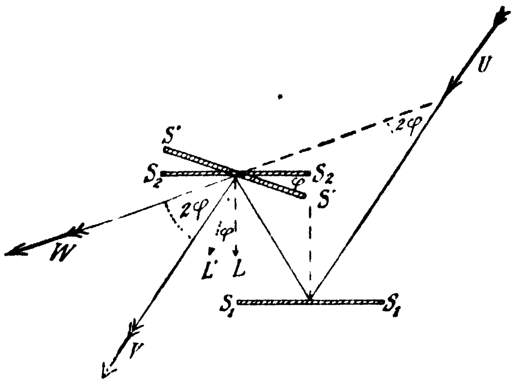


Fig. 6.—To illustrate the principle of the sextant

Mirrors inclined at an angle to one another.—The two mirrors S_1S_1 and S_2S_2 (fig. 6) are set up parallel to one another. The ray of light U falls on S_1S_1, is reflected, falls on S_2S_2, and is reflected in the direction V. As the angles of incidence of the ray are the same for the two parallel mirrors, and hence the angles of reflection are equal, the twice-reflected ray V is parallel to the incident ray U. We now rotate the mirror S_2S_2 through the angle ϕ into the position $S'S'$. The normal to the mirror (L) is thereby also rotated through the angle ϕ into the position L′, and the angle of incidence of the ray falling on $S'S'$ is increased by ϕ; hence the angle of reflection is increased by the same amount. Consequently the reflected ray is rotated through the angle 2ϕ. Hence it follows that the new reflected ray W now makes the angle 2ϕ with the original ray U.

If a ray of light is reflected in succession by two mirrors inclined at

an angle to one another, the ray is deviated through an angle which is equal to twice the angle between the mirrors.

This fact is utilized in the **sextant** (fig. 7). A frame in the form of a sector of a circle (of about 60°, hence the name sextant) is graduated along the circular edge MC. A small plane mirror B whose plane is at right angles to that of the sector is free to rotate about its axis, which passes through the centre of the circle. The mirror is fixed to one end of the arm *Bn*; the other end of the arm carries a vernier which enables the position of the arm to be read off accurately on the graduated circle. Facing the rotating mirror there is a second mirror A, which is fixed to one limb of the sector and has only its lower half silvered. When the plane of the rotating mirror B is parallel to the plane of the fixed mirror A the zero point of the movable arm must coincide with the zero of the divided circle. If we look through the hole O at the silvered and non-silvered halves of A simultaneously, two pencils of rays enter the eye at the same time, namely,

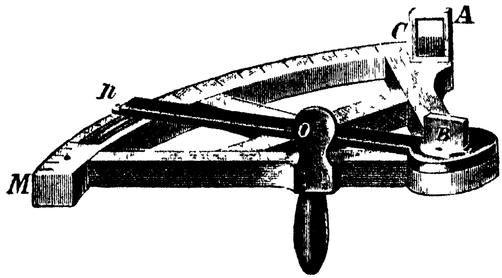


Fig. 7 —Model of a sextant

the one coming directly through the non-silvered part of A and the one reflected from B to A and then from A to O by the silvered half of A. Thus, e.g., if we look through O at a star while the two mirrors are parallel to one another we see the star once through the glass A and (if the sextant is slightly tilted) again close beside it in the mirror part of A. If the mirror B is now rotated, a ray which is to be reflected to A and thence to the eye through O must fall on B from another direction, e.g. from another star. Hence we then see two stars, one on top of the other, the angle subtended by them at the eye being twice the angle through which B was rotated. The latter angle can be read off on the divided circle, the graduations of which give the angle subtended by the two stars directly, i.e. give twice the angle through which the mirror has actually rotated.

Fig. 7 shows a very simple model sextant with a wooden frame. For actual observations the apparatus is made of metal and the graduations are made on a silvered strip of metal. The hole O is replaced by a telescope, which enables the adjustment to be made very accurately. With the sextants in ordinary use angles can be measured accurately to within half a minute of arc.

The sextant is a very convenient instrument for measuring the angle subtended at the eye by two stars, the angular elevation of the sun, and so on. The chief advantage of the sextant is that it requires no fixed support; hence its extensive use in navigation.

The **Kaleidoscope.*** In the kaleidoscope the repeated reflection of an object at two mirrors meeting one another at an acute angle is used to produce a multiplicity of patterns, all of which are symmetrical about the centre.

* Gr., *kalos*, beautiful; *eidos*, form; *skopein*, to see.

3. Curved Mirrors.

Concave Mirrors.—Some of the properties of concave mirrors have already been deduced from the general theory of wave motion (Vol. II, pp. 217, 246). It will be instructive for the reader to renew his acquaintance with them, starting from the point of view of geometrical optics.

Parallel Rays

The laws of the plane mirror may also be applied to small elements of curved surfaces. The normal to such an element is the normal at that point to the surface.

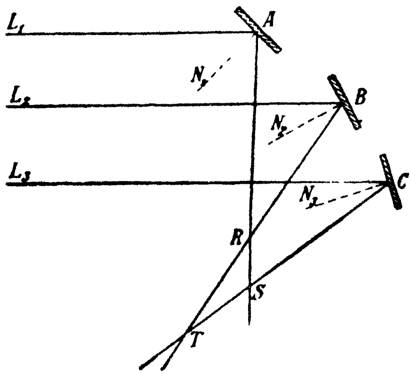


Fig. 8.—Surface elements of a concave mirror

Suppose that L_1A and L_2B in fig. 8 are two rays belonging to a parallel pencil, which are very close to one another and are incident on the two neighbouring surface elements A and B of a curved mirror. Let AN_1, BN_2 be the normals at A, B to the surface of the mirror. The rays reflected by the surface elements intersect at R. According to the law of reflection $\angle L_1AN_1 = \angle N_1AR$ and $\angle L_2BN_2 = \angle N_2BR$. The ray reflected by a neighbouring surface element C in general intersects the two first reflected rays at two points S and T which do not coincide with R. In the special case where the surface elements A, B, C are parts of a paraboloid of revolution with its axis parallel to the incident rays, the three points of intersection R, S, and T coincide with the geometrical focus of the paraboloid of revolution, for the focal radius and the diameter of a parabola are equally inclined to the normal of any point (fig. 9). Hence:

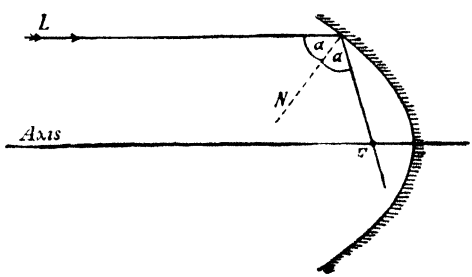


Fig. 9.—Parabolic mirror

If a parallel pencil of rays falls on a mirror with the shape of a paraboloid of revolution in the direction parallel to its axis, the reflected rays all pass through the so-called focus of the parabolic mirror (fig. 10).

From the principle of reversibility of the path of the rays we have the following result:

The rays proceeding from a point source of light at the focus of a parabolic mirror, and in general any rays passing through the focus, are reflected in a direction parallel to the axis of the mirror.

These relationships have already been used in the discussion of Hertz's experiments (Vol. III, p. 637).

The radius of the circle of curvature at the vertex of a parabola is twice the focal length, so that M, the centre of the circle of curvature (fig. 11) is twice as far from the vertex S as the focus F. This circle of curvature has especially close (four-point) contact with the parabola, so that near the vertex it may be substituted for the parabola. Hence it follows that we can replace a *parabolic* mirror by a *spherical* mirror provided we use only the central part of the mirror and the radius of the sphere is twice the focal length of the parabolic mirror.

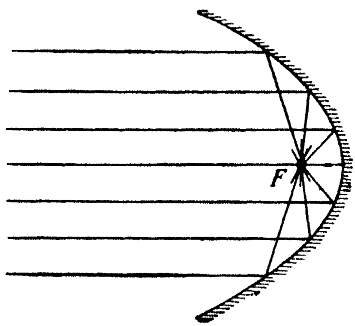
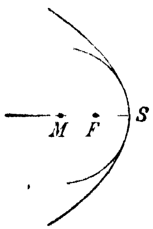
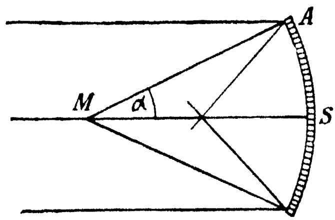


Fig. 10.—Parallel rays incident on concave mirror

A **concave mirror** consisting of part of a spherical shell is called a **spherical mirror;** the straight line which passes through the centre of curvature M (fig. 12) and the centre of the surface or **pole** of the mirror (S) is called the **principal axis,** or simply the **axis,** of the mirror. Twice the angle SMA (α) formed by the axis MS and the radius MA drawn to the outer edge A is called the **angular aperture** * of the spherical mirror.

If rays fall on a spherical mirror in a direction parallel to the axis, after reflection they all pass through one point, which is equidistant from the centre of curvature and the pole of the mirror.





Figs. 11, 12.—Parallel rays incident on concave mirrors

This point is called the **principal focus** of the mirror.

The distance of the principal focus from the pole is called **the focal length** (f). If the radius of curvature of the mirror is r, we have

$$f = \tfrac{1}{2} r.$$

If the angular aperture is large, the spherical mirror does not reflect all the rays parallel to the axis in such a way that they meet in a singl

* [The diameter of the circular boundary of the mirror is sometimes called the *aperture*.]

point. The points where neighbouring rays intersect after reflection lie on a surface (**caustic * surface**) which has a cusp at the principal focus of the mirror.

The caustic surface becomes visible if sunlight is allowed to fall on the concave surface of a silvered clock-glass and smoke is blown into the path of the rays; it appears bright as compared with the surrounding space. The peculiarly-shaped bright line seen when sunlight shines on a polished napkin-ring or finger-ring lying on a white background is the intersection of the caustic surface with the plane of the background, the **caustic curve** (fig. 13, Plate III).

Divergent and Convergent Rays

Object and Image on the Axis.—The two focal radii of a point on an ellipse make equal angles with the normal at that point. Hence if a source of light is placed at one focus of a mirror which is an ellipsoid of revolution, the rays are reflected so as to pass through the other focus. The two points determine the axis of the ellipsoid.

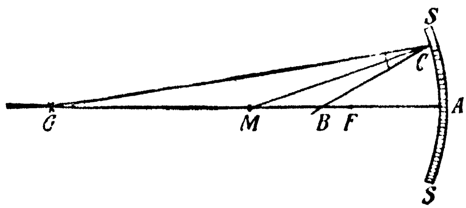


Fig. 14.—Divergent rays incident on a concave mirror

Provided the angular aperture is small, an elliptical mirror may likewise be replaced by a spherical mirror whose radius is equal to the radius of the circle of curvature at the vertex of the ellipse.

It is therefore possible to produce a mirror of spherical curvature which will cause rays parallel to the axis and also divergent rays proceeding from a definite point on the axis to meet in one point, provided only that the angular aperture of the mirror is sufficiently small.

Let SAS (fig. 14) be a spherical mirror with pole A and centre of curvature M. Let G be a point source from which divergent rays fall on the mirror, such as GC and GA. For the first of these CM is the normal to the surface; GA is reflected back along itself. B, as will be shown, is the point where all the rays emitted by G and reflected by the mirror meet. It is called the *image* of the *object* G. The ray GC is reflected to B in such a way that $\angle GCM = \angle MCB$. As CM is the bisector of the vertical angle in the triangle GCB, we have

$$GC : CB = GM : BM.$$

Convention as to Sign.—Here we shall lay down the following convention about the signs of the quantities involved in calculations connected with reflection. *The light is always to travel from left to right.*

* Gr., *kaustikos*, burning.

The distance from a mirror of a point emitting or receiving light is always to be reckoned positive in the direction in which the light is travelling.

GC and GA are accordingly positive in fig. 14, as the light is moving from G to the mirror. BC and BA are negative, as the light is moving from the mirror to the point B. We shall, however, make the origin from which GA and BA are reckoned at A, the pole of the mirror, that is, the intervals are to be reckoned in the direction from A to G and from A to B once and for all. We make this clear by calling them AG and AB (in contradistinction to GA and BA). Then AG in fig. 14 is negative, AB positive, for the direction AG is opposite to the corresponding direction of motion of the light, while the direction AB is in the same direction. That is, we are to take AG = − GA and BA = − AB.

The distance of A, the pole, from M, the centre of curvature of a mirror, is always to be taken positive when M lies to the *left* of the pole (the reflected ray moving towards the centre M).

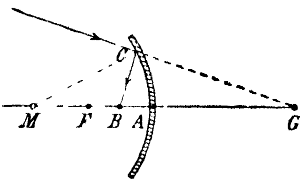


Fig. 15.—Convergent rays incident on a concave mirror

Fig. 15 shows the course of the rays when AG and AB have positive numerical values. The object G on the axis of the mirror now lies behind the mirror: it does not now emit divergent rays, but rays converge towards it. The object in this particular case is **virtual** (p. 6), as it is only the *prolongations* of the incident rays that meet there. The corresponding image B is **real.**

The Object-Image Equation.—In fig. 15 CM is the bisector of the exterior angle of the triangle GBC. Hence, as before, we have

$$GC : CB = GM : BM.$$

For brevity we put AG = u, AB = v, and AM = $2f$. (According to the conventions given above, therefore, u is negative in fig. 14, positive in fig. 15; v and f are positive in both.) AF = $\frac{1}{2}$AM = f is called the **focal length.**

The concave mirror has a positive focal length f*: the object is virtual if* u *is positive, real if* u *is negative: the image is real if* v *is positive, virtual if* v *is negative* (see above).

Restriction to Rays in the Neighbourhood of the Axis.—For small angular apertures C falls in the neighbourhood of A and the rays are inclined at a very small angle CGA to the axis. Such rays are called **axial rays.** For such rays GC may be replaced by GA and CB by AB. With this limitation the statement of proportion given above may be written

$$GA : AB = GM : BM.$$

Referring to fig. 15, we may replace this by

$$-u : v = -(2f + u) : (2f - v),$$

which on rearrangement becomes the equation

$$\frac{1}{v} = \frac{1}{u} + \frac{1}{f}$$

giving the relationship between the positions of object and image:

The reciprocal of the distance of the image from the pole is equal to the sum of the reciprocals of the distance of the object from the pole and of the focal length.

As according to this equation the value of v is independent of the particular ray selected, the equation tells us that with the above limitation all the reflected rays are united at B.

Note.—Very commonly the distance of the object from the pole of the mirror is reckoned positive in the direction opposite to that of our convention. The equation then becomes symmetrical with respect to the quantities u and v. Our convention was adopted so as to be consistent with that which we shall use later in discussing the formation of images by lenses.

The Correspondence between Image and Object.—To every object-point G there corresponds an image-point B, and conversely. B and G are called **corresponding points, or conjugate foci.** The correspondence is such that B and G can exchange places, i.e. if the object moves to the point where its image was, the image moves to the point where the object was (involutory * correspondence).

If, in particular, $u = -2f$, i.e. the object is at the centre of curvature, $v = 2f$. That is, object and image meet at the centre of curvature. The same is true for the pole, where $u = v = 0$.

If $u = \infty$, $v = f$: similarly if $u = -f$, $v = \infty$. That is:

The focus is the image of a point at infinity; all rays falling on the mirror parallel to the axis are therefore united at the focus.

The image of the focus is the point at infinity on the axis; rays proceeding from the focus and falling on the mirror are therefore reflected parallel to the axis.

Newton's Equation connecting the Focal Distances.—Instead of reckoning all distances from the pole A, we may use the distances of the object and image from the focus.

We shall call the distances $|FG| = |g|$ and $|FB| = |b|$ (fig. 15) the **focal distances** of the object and image, and we shall reckon g positive to the right of the focus (like u), b positive to the left (like v). Then we have †

$$g = FG = f + u,$$
$$b = FB = -f + v,$$

or

$$u = g - f, \quad v = b + f.$$

* Lat., *involvere*, to turn round; involutory accordingly means something that can be reversed.

† $|b| = |FA| - |BA| = f - v$; and b is negative in fig. 15.

The equation giving the relationship between the positions of object and image (see above) then becomes

$$\frac{1}{b+f}=\frac{1}{g-f}+\frac{1}{f}.$$

On rearrangement this gives *Newton's equation for the focal distances*

$$gb=-f^2.$$

Equation of Convergence. The relationship between the positions of object and image may also be deduced from a different point of view, which enables the facts to be expressed in another way which has a variety of advantages. According to fig. 14 $\angle$CMA is an exterior angle of the triangle CGM. Hence

$$\angle CMA = \angle CGA + \angle GCM.$$

As $\angle$CBA is an exterior angle of the triangle CMB, we similarly have

$$\angle CMA = \angle CBA - \angle BCM.$$

Adding the equations and taking account of the equality of the angles at C,

$$2\ \angle CMA = \angle CGA + \angle CBA.$$

If we now imagine arcs drawn with centres G, B and radii GC, BC, these will meet the axis nearer A the smaller the angular aperture of the mirror. For *small angular apertures,* these two arcs may therefore be taken equal to the arc AC If we now express the angles of the last equation above in radian measure, bearing in mind that in fig. 14 GA and BA have opposite signs, the angles are given by the ratios of the corresponding arcs and their radii. We thus obtain

$$2\frac{AC}{MA}=\frac{AC}{GA}+\frac{AC}{BA}$$

or

$$\frac{1}{v}=\frac{1}{u}+\frac{2}{r}=\frac{1}{u}+\frac{1}{f}.$$

From the above argument we see that $1/u$ is a measure of the magnitude of the angle CGA or a measure of the angular aperture of the *pencil of rays* falling on the mirror from G, if we regard C as the outermost part of the mirror. Similarly, $1/v$ is a measure of the angular aperture of the pencil of rays which is reflected back by the mirror so as to converge at B. Hence the quantities $U=1/u$, $V=1/v$ may be called the *convergences* of the object-point and image-point respectively. If, further, we call $1/f=$ D the **power of the mirror,** we obtain

$$\mathbf{V}=\mathbf{U}+\mathbf{D},$$

a simple equation due to Gullstrand.* That is:

The convergence of the image-point is always equal to the sum of the convergence of the object-point and the power of the concave mirror.

The Relationship between the Distances from the Pole and the Focal Distances.

According to our convention (p. 40) the object G (fig. 14) is real

* A. Gullstrand, Professor of Physiological Optics at Upsala.

when it lies to the left of the concave mirror, i.e. when $-\infty \leq u \leq 0$. Now $1/v = 1/f + 1/u$ is also positive, i.e. the image B is real, provided $-\infty \leq u \leq -f$. When $u = -f$ a change of sign occurs in v, which becomes negative, i.e. the image is on the right of the mirror and therefore *virtual*, when $-f < u < 0$. That is, if the real object moves to the right from infinity on the left towards the mirror, the real image moves to the left from the focus F away from the mirror. The two meet at the centre of curvature. When the object reaches the focus, the image has moved to an infinite distance on the left. If, however, the real object now moves into the region FA, a *virtual* image appears to the right of A and moves towards the object. The two points meet again at A, where $u = v = 0$. Thereafter all virtual object-points to the right of A give rise to real image-points between A and F

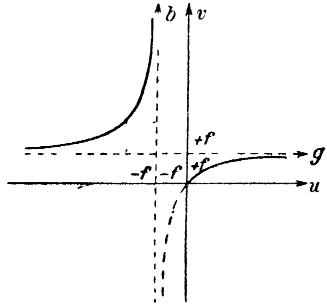


Fig. 16.—Graphical representation of the relationship between the positions of object and image, u, v, are the distances from the pole, b, g the focal distances

The relationship $1/v = 1/u + 1/f$ between the positions of the object and the image is illustrated graphically in fig. 16. The part of the curve corresponding to real images is continuous, the remainder dotted. The curve is an equilateral hyperbola. If the *focal distances* b and g are taken as variables the axes (dotted in fig. 16) are the asymptotes of the hyperbola.

Properties of the Image

Subsidiary Axes.—Any straight line drawn through the centre of curvature of a concave spherical mirror may be regarded as an axis.

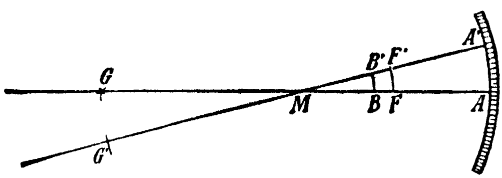


Fig. 17—Principal axis and a subsidiary axis of a mirror

In contradistinction to the *principal* axis introduced above, which joins the pole of the mirror to the centre of curvature, we shall call such a line a **subsidiary axis.** Then if we draw a straight line through the point A′ at one side of A (fig. 17) and the centre of curvature M and describe circles with centre M passing through F, B, and G and cutting the subsidiary axis A′M at F′, B′, and G′, the image B′ corresponds to the object G′.

Plane Image.—Provided the angular aperture is small, we may replace the arcs FF′, BB′, GG′ by the perpendiculars at F, B, G. Hence if the object at G consists of a *small* plane surface at right angles to the principal axis, it gives rise to a plane image at B similar to the object, perpendicular to the principal axis, and upside down.

Magnification.—The ratio BB′ : GG′ = m is called the **magnification** (or *magnifying power*)

By fig. 17 we have BB′ : GG′ = MB : MG. Hence $m = \frac{MB}{MG} = \frac{AB - AM}{AG - AM}$. Now $AB = v$, $AM = r = 2f$, and $AG = -u$ (p. 40). Hence $m = \frac{v - 2f}{-u - 2f}$. Using the equation $\frac{1}{v} = \frac{1}{u} + \frac{1}{f}$ (p. 41) or $f = \frac{uv}{u - v}$, it follows that $m = \frac{v}{u}$, as may also be deduced immediately from fig. 17.

Construction of the Image.—Fig. 18 shows how the image A′ of a point A not lying on the axis may be found graphically, provided the centre of curvature M and the focus F of the mirror are known. From A we draw the ray An parallel to the principal axis; this is reflected so as to pass through the focus. Further, we draw the ray AM through

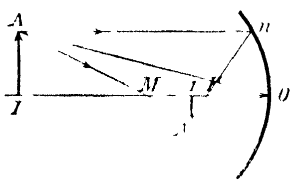


Fig. 18.—Construction of the image

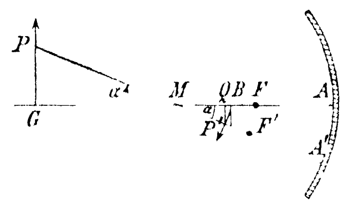


Fig. 19.—Curvature of the image formed by a concave mirror

the centre of curvature; this is reflected back into itself. Then the intersection A′ is the image of A; all other rays proceeding from A intersect at this point.

The point O (fig. 18) where the principal axis MF meets the mirror is also called the **optical centre** of the mirror. The ray AO is called the **central ray** of the pencil of rays proceeding from a point A and falling on the mirror.

Curvature of the Image.—It is only when the pencil of rays proceeding from the object is of small angular aperture that the rays proceeding from one point of the object are reunited with a sufficient degree of accuracy in a single image-point (p. 39). Hence it is only small plane elements of surface at right angles to the axis that are represented by plane images. We shall now consider a special case where any point of the object is represented by a definite point even when the rays are inclined to the axis at a fairly large angle, but the image appears curved.

Let M (fig. 19) be the centre of curvature, A the optical centre, and F the focus of a concave mirror, and suppose that at G there is a plane luminous object at right angles to the principal axis MA, and that a pencil of rays starts from the point P of the object. *We shall assume that this pencil is of small angular aperture and that its central ray passes through the centre of curvature* M. Let B be the image of G, the point of the object which is on the axis; then P′, the image of the point P, is on the subsidiary axis A′P. By Newton's equation on p. 42, putting $|FB| = |F'P'| = |b|$ and $|FG| = |F'P| = |g|$ (where, as on p. 41, b is positive and g negative), we have

$$f^2 = F'P \,.\, F'P' = (MP + f)(f - P'M).$$

Hence $P'M = MP \,.\, f/(MP + f)$. Then if the subsidiary axis is inclined at the angle α to the principal axis, $MP = \dfrac{FG - FM}{\cos\alpha} = \dfrac{-(g+f)}{\cos\alpha}$, and MQ, the projection of MP′ on the principal axis, is $MP' \cos\alpha$. On substituting and rearranging, we have

$$QM = \frac{(g+f)\, f \cos\alpha}{g + 2f \sin^2 \alpha/2}.$$

If through B we describe a plane at right angles to the principal axis GA, the image-point P′ is at the distance $QB = MB - MQ$ from this plane. Here $MB = MF - BF = -f + b = -f - f^2/g$. Substituting for MB, MQ, and rearranging, we obtain

$$BQ = -QB = -\frac{2f \sin^2\frac{\alpha}{2}(g+f)^2}{g\left(g + 2f \sin^2\frac{\alpha}{2}\right)}.$$

The value of BQ is accordingly very small when the angle of inclination α is small, but increases rapidly as α increases. Hence in this particular case:

The farther a point of a plane object at right angles to the axis is from the axis the farther removed is its image from the image-plane determined by the point of the object which is on the axis.

The image is curved, the convex side being towards the pole of the mirror. Hence if a plane screen placed at B is used to receive the real image, only the central part of the image will be clearly defined, the outside parts appearing blurred. For the point where the rays producing a lateral point of the image intersect lies farther from the mirror, so that the rays have not yet met by the time they reach the screen and the intersection of the cone of radiation with the plane screen is a small circle. These circles are called **circles of confusion.**

Defects in the Image. Stops.—The indistinctness due to the circles of confusion is called **spherical aberration.** In general, defects in the image are the result of defects of symmetry, as is shown, e.g., on comparison of the paths of the rays in reflection at a parabolic mirror with those in reflection at a spherical mirror. They are specially evident when the system has a large aperture or the object is large. As in the case of the spherical mirror there is no possibility of correcting

the aberration by other means, it is necessary to confine ourselves to small apertures and small angles.

In optical instruments involving the use of concave mirrors, therefore, only those of *small angular aperture* may be employed. To avoid the use of the light rays falling on the margin of a large mirror and the resulting indistinctness, opaque screens with circular holes, called **stops** or **diaphragms**, are placed in the path of the rays. The action of stops is discussed in detail in the next section.

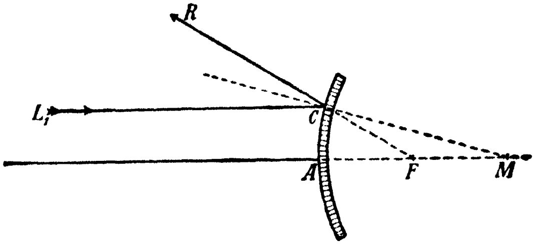


Fig. 20.—Virtual focus of a convex mirror

Convex Mirrors.—In the curved mirrors discussed hitherto the *concave* side was used as reflecting surface and the mirror was accordingly referred to as a *concave* spherical mirror. If the *convex* side is made the reflecting surface, the mirror is called a *convex* spherical mirror. From fig. 20, which represents a convex mirror with centre of curvature M, we see that a ray L_1C parallel to the axis is reflected so as to *diverge* away from the principal axis. The backward prolonga-

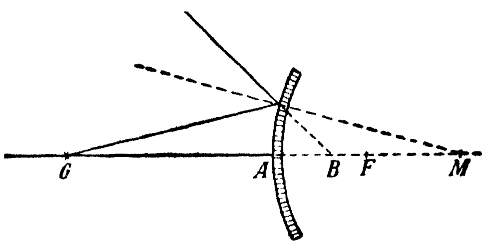


Fig. 21.—Convex mirror giving rise to a virtual image of a real object

tion of the reflected ray CR intersects the principal axis produced at F. Here again F is midway between A and M. We call F the **virtual focus** and AF the **virtual focal length**; as f is reckoned in the direction opposite to that for the concave mirror, it may also be referred to as the negative focal length.

In the case of the convex mirror, real objects can give rise only to virtual images, for the rays from the point source G (fig. 21) always diverge after reflection by the mirror. The backward prolongations of the rays intersect at a point B behind the mirror, the virtual image; by p. 40, its distance from the pole (AB) is to be reckoned negative.

The equation giving the relationship between the distances of object and image from the pole and the focal length (deduced as in the case of fig. 15) is again

$$\frac{1}{v} = \frac{1}{u} + \frac{1}{f}.$$

In this equation the *focal length* is accordingly to be taken *negative.*

The construction of the image for points not on the axis of the convex mirror is carried out in fig. 22. A ray An parallel to the axis is reflected in such a way that F, the virtual focus of the convex mirror, is the apparent starting-point of the reflected ray nv. The ray AM in

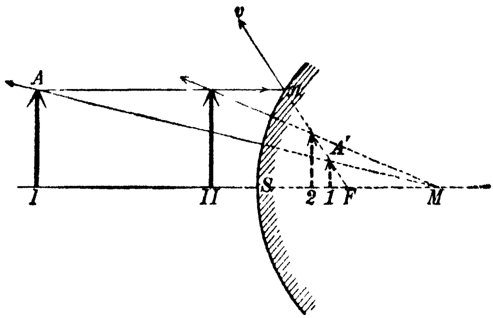


Fig. 22.—Construction of the erect, diminished, and virtual image produced by a convex mirror

the direction of the centre of curvature is reflected back into itself. Then A′ is the apparent starting-point of the reflected rays, i.e. the virtual image of A. In the figure the construction is shown for two different positions I and II of the same object.

If the object is real and erect the image produced by the convex mirror is virtual, erect, and diminished in size.

4. Stops or Diaphragms.

Out of all the rays which are emitted by one point of a source of light the pupil of the eye selects only a very small portion (fig. 23) and these are alone concerned in vision. The central ray of the pencil entering the eye is the ray from the source to the centre of the pupil. The larger the pupil the greater is the radiant energy entering the eye and the brighter does the source of light appear. The brightness also depends on the distance of the source from the eye. For as the rays spread out in straight lines, the area of the cross-section of a limited pencil of rays is greater the greater the distance from its vertex. Hence if two eyes at different distances from a point source are to receive equal amounts of light (equal luminous flux), the diameters of the pupils must be proportional to the distances from the source of light.

It follows from this that the brightness of the image actually perceived depends not only on the intensity I of the source of light but on the size of the cone of light whose vertex is the point source and whose base is the opening of the pupil. This cone is called the

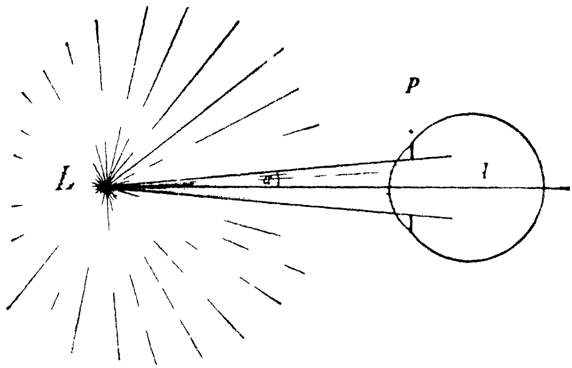


Fig. 23.—The pupil of the eye as an aperture stop

aperture cone and the (small) vertical angle of the cone ($2a$) is called the (angular) **aperture.** The base of the cone (the space angle) is proportional to a^2, hence the brightness of the image of a luminous point is determined by the expression Ia^2.

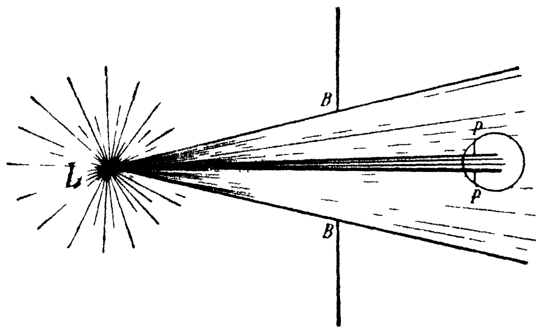


Fig. 24.—A field stop BB

As the pupil stops out from the whole collection of rays the cone of rays which determines the aperture, it may be called an **aperture stop.***

It may happen that there is also another stop present which selects a portion from the whole series of light rays. If this stop is arranged like BB in fig. 24, it has no effect on the aperture of the cone of rays

* Ger., *Aperturblende*

entering the eye. If, however, we imagine the source of light L moved upwards parallel to the stop BB, the position relative to the eye of the cone of light passing through the stop BB is altered; its upper boundary moves nearer and nearer to the pupil and may pass the pupil by altogether. The source of light is then no longer visible to the eye. The totality of all the points of space from which light can reach the eye, i.e. the region visible to the eye, is called the **field of view.** The stop BB restricts the field of view, so that we may call BB a **field stop.***

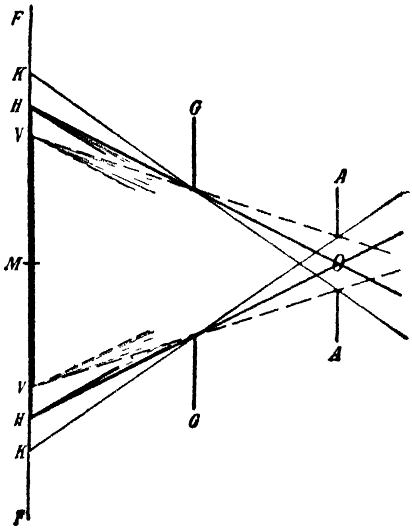


Fig. 25.—An aperture stop AA, and a field stop GG

In fig. 25 FF denotes a plane of which all points are sources of light, e.g. an illuminated (i.e. light-emitting) picture hanging on the wall. The stop GG is placed at an arbitrary distance from it, e.g. it may be a window in the opposite wall of the room. Let AA be an aperture stop, e.g. the pupil (drawn on an exaggerated scale) of an eye at rest looking at M.

The point M behaves like the point L of the previous figure. The stop GG has no effect on the cone of rays entering the aperture stop AA from M. Points in the neighbourhood of M behave in a similar way. The outermost points for which GG has no effect are the points VV, as they still send out a complete cone of light (indicated in the figure), which fills the whole pupil AA.

The points adjacent to VV on the outside send to the aperture stop AA cones of light whose rays are only partly allowed through by the stop GG. The light rays emitted by HH only half fill the stop AA, and the outer boundary ray still passes through the middle of AA. Finally, the points KK send cones of light through the stop GG, none of which enter the opening of the aperture stop AA. All the points lying outside KK behave in a similar way.

We see that GG bounds the field of view (of an immovable eye with pupil AA at O), so that GG is a field stop. The field of view is not sharply bounded off; on the contrary, the limitation begins at VV and ceases at KK. When we use the idea of a field of view in future, we shall mean the part of the field of view limited by points which send rays into the centre of an eye at rest. These are the **principal rays.** In this sense the field of view is the whole surface lying between HH.

* Ger., *Gesichtsfeldblende.*

Note the distinction between principal rays and the central rays of image-forming pencils of rays (p. 44); the position of the latter is determined by the point source and the centre of that aperture stop (lens holder or diaphragm opening) that limits the course of the rays in the pencil most.

Summarizing, *the brightness of the image seen depends on the aperture stop.*

The size of the field of view is determined by the rays which are drawn from the centre of the aperture stop to the boundary of the field stop.

Note.—If stops are used in an instrument containing any optical parts (mirrors, lenses, &c.), these act partly as field stops, partly as aperture stops, in so far as the pupil of the eye does not play the part of one or the other. The outer edges and holders of lenses and mirrors may also act as stops. We have first to find out which stop is acting as an aperture stop; this is in general the one which appears to subtend the smallest angle at the object. Either the stop itself, if it lies between the object and the first optical part of the instrument, or its optical image, at the centre of which the vertex of the cone of light limiting the field of view is situated, is the so-called **entrance pupil** * of the instrument. This is the base of any image-forming cone of rays starting from the points of the field of view.

The entrance pupil is the aperture stop of the whole optical system.

If the eye of the observer is moved to the centre of the entrance pupil, the other (concentric) stops or stop-images appear as concentric circular windows through which he observes the object. The smallest of these stop-images limits the field of view from the object side; it may be called the **entrance window.**†

The entrance window is the field stop on the object side.

5. The Refraction of Light.

As was deduced from Huygens' principle in Vol. II (p. 248), every wave motion is subject to refraction at the boundary between two media. We shall now investigate these phenomena in more detail for the case of light.

Results of Observation.—A coin M is placed on the bottom of an opaque dish and the observer stands in such a position that the coin is just hidden by the top edge of the dish. If water is then poured into the dish, the penny becomes visible, although the observer's eye is still in the same place; it appears to be raised to M_1 (fig. 26) by the water.‡

In order to trace out the behaviour of a light ray entering water, we make the following experiment. A vessel partly filled with water has a white disc, on which a circle is drawn, fixed to its rear wall (fig. 27). The disc is dipped so far into the water that JK, the diameter of the circle, lies in the surface of the water. LR is at right angles to the surface of the water. If a narrow pencil of light E is let fall obliquely on the surface of the water in such a way that it meets the boundary between air and water at the centre of the circle M, the pencil splits up into two parts. One is reflected back in the direction MS, the other enters the water. The pencil entering the water, however, does not proceed in the

* Ger., *Eintrittspupille.*

† Ger., *Eintrittsluke.*

‡ This experiment was known to Archimedes (died 212 B.C.).

original direction, but is shifted nearer to the diameter LR. If the direction of the incident pencil is altered, the direction of the pencil traversing the water is always altered. The customary notation for the phenomena of refraction is given in fig. 28.

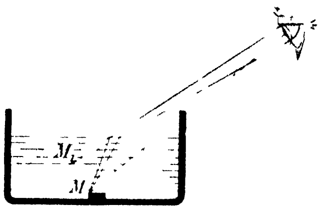


Fig. 26.—Apparent raising of an object as a result of refraction

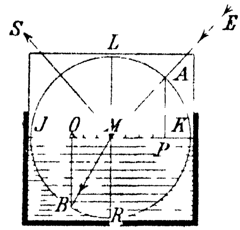


Fig. 27.—Refraction and reflection

Reversibility of the Path of Light.—If the refracted ray of light entering the water at O is allowed to fall perpendicularly on a plane mirror in the water so that the ray is reflected back into itself, this ray of light emerging at O is subject to a refraction away from the normal and of such a magnitude that it exactly coincides with the incident ray in the air also, i.e. returns into itself in the air also.

If the direction of a ray of light is reversed, the light retraces its former path both in reflection and refraction; or, in other words, *any path which is such that light can traverse it can be traversed by light in either direction* (**reversibility of the path of light**).

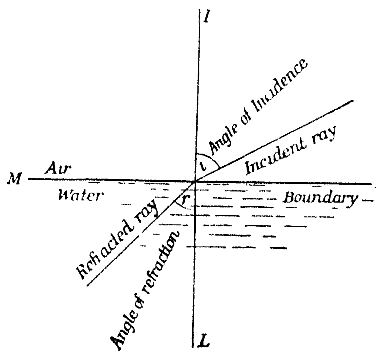


Fig. 28

This theorem, which we assumed in advance on p. 32, is only valid for geometrical optics under the assumptions and simplifications there mentioned. In view of this reversibility, we shall sometimes refer to both the angles i and r in fig. 28 as *angles of refraction* without distinguishing between them.

Refractive Index.—From suitable experiments we learn that:

(1) *If a ray of light passes from air into another transparent body, its deviation from the original direction is greater the more obliquely it strikes the boundary surface.*

(2) *The incident ray, the normal at the point of incidence, and the refracted ray all lie in one plane.*

If the angles of refraction are measured for a number of rays falling on the plane boundary of a body in different directions, it is found that there is a definite relationship between the two angles.

In fig. 27 AM is the incident ray, MB the refracted ray. If from the point A where the incident ray cuts the circle we draw the perpendicular AP to the boundary surface JK and from the point B where the refracted ray cuts the circle we draw the perpendicular BQ to JK, we find that

MP *and* MQ *are always in the same numerical ratio, no matter what the magnitude of the angle of incidence is.*

When a ray of light passes from air to water, this ratio MP/MQ is equal to 4/3. The ratio 4/3 is called the **index of refraction** or **refractive index** of water relative to air (${}_a\mu_w$).

If, however, we take the radius of the circle as unit, MP is the sine of the angle MAK, i.e. is equal to $\sin i$ (fig. 28); similarly, MQ is equal to $\sin r$. We have the relationship

$$\frac{\mathrm{MP}}{\mathrm{MQ}} = \frac{\sin i}{\sin r} = \text{const.} = {}_a\mu_w.$$

*The ratio of the sines of the angles of refraction for two given materials is constant.**

Absolute Refractive Index.—In ordinary circumstances the ray of light is allowed to pass from air to some other transparent body, such as water or glass. If the ray of light passes from a *vacuum* into a transparent body such as glass, the ratio of the sines of the angles of refraction differs somewhat from the value obtained when the ray of light passes into the body from air. The refractive index for the passage from vacuum to transparent body is termed the **absolute refractive index** or briefly the refractive index.

Note.—The value obtained in air is often quoted as the refractive index.

The Law of Refraction.—It was shown in Vol. II (p. 249) that for any wave motion refraction is a perfectly simple consequence of the varying rates of propagation of the waves. We obtained the law which we have just deduced experimentally in the case of light, namely, that the quotient of the sines of the angles of refraction is constant and equal to the ratio of the velocities in the media under consideration. We have

$$\frac{\sin i}{\sin i'} = \frac{v}{v'} = \mu,$$

where v and i are the velocity and the angle of refraction in one medium,

* This relationship was first discovered by SNELL (WILLEBRORD SNELL VAN ROYEN, latinized as Snellius, a Dutchman (1581–1626)) in 1618; it was, however, not widely known until DESCARTES published it in his *Dioptric*, probably independently of Snell, in 1637. The first attempts to deduce a law of refraction go back to CLAUDIUS PTOLEMÆUS of Alexandria (150 B.C.). His measurements of the angles of refraction at water and glass have come down to us and probably represent the most ancient physical experiments recorded historically.

v' and i' the corresponding quantities in the other. We may write this in the form

$$\frac{\sin i}{v} = \frac{\sin i'}{v'} = \text{const.}$$

As the paths of the light-rays are reversible, by the experiment described above, it makes no difference here whether the ray passes from medium 1 to medium 2, or from medium 2 to medium 1. If one medium is a vacuum, $v = c = 3 \cdot 10^{10}$ cm./sec. We shall transform the above equation by multiplying it by c:

$$c\,\frac{\sin i}{v} = c\,\frac{\sin i'}{v'};$$

but c/v is the ratio of the velocity of light in a vacuum to that in the medium under consideration, i.e. is equal to the absolute refractive index (cf. Vol. II, p. 249). If we denote the refractive index of one medium by μ and that of the other by μ', we have the *law of refraction*,

$$\boldsymbol{\mu \sin i = \mu' \sin i' = \text{const.}}$$

That is:

The product of the index of refraction and the sine of the angle of refraction is constant.

The product $\mu \sin i$ is called the **numerical aperture** * of the ray relative to the normal at the point of incidence. We may therefore say:

In refraction the numerical aperture relative to the normal at the point of incidence remains constant.

Thus the path of a light ray through the most varied media (provided they are in parallel layers) may be conveniently traced merely by using the fact that the numerical aperture is constant. In this case, therefore, the numerical aperture is an *optical invariant.*

The value 4/3 found by experiment (p. 52) represents the index of refraction of water relative to air ($_a\mu_w$). From the law of refraction we have

$$\sin i_a \cdot \mu_a = \sin i_w \cdot \mu_w,$$

so that

$$_a\mu_w = \frac{\mu_w}{\mu_a}.$$

The refractive index of one medium (e.g. water) relative to another

* This term was introduced by E. Abbe (1840–1905), professor at Jena, one of the founders and subsequently the sole partner of the Carl Zeiss optical works. Not only did he make important contributions to both theoretical and practical optics, but he was also distinguished by his social work, and unselfishly devoted his means to the encouragement of social and industrial reforms.

(e.g. air) is equal to the ratio of the refractive indices of the two media relative to a third (e.g. a vacuum).

This enables us to transform refractive indices measured relative to air into refractive indices relative to a vacuum. The refractive index of air (at 20° C.) relative to a vacuum is $\mu_0 = 1{\cdot}00028$. Hence if the refractive index of a substance relative to air is μ and that relative to a vacuum is μ', we must have

$$\mu = \frac{\mu'}{\mu_0} \text{ or } \mu' = \mu\mu_0 = 1{\cdot}00028\mu.$$

Note.—We do not see an object unless it is self-luminous or reflects light. In the latter case the sum of the energies of the refracted and reflected light is equal to the energy of the incident light. Thus if a ray of light falls on the boundary of two media which have the *same* refractive index, there is no refraction; the ray of light passes through without change. Hence no light is reflected at all, i.e. the object is invisible; for example, a thick block of glass immersed in cedarwood oil cannot be seen. Again, a method has recently been devised for making animal tissues which have been bleached by hydrogen peroxide perfectly transparent; the tissue is carefully freed of water and embedded in oil of wintergreen, which has the same refractive index as the tissue. The bones, or the circulatory system when filled with mercury, can be observed within the undamaged tissue. For the same reason the air is not as a rule visible, as the rays of light emitted by a source reach the eye without change of direction. On the other hand, we can see air in the form of bubbles in water or glass, because refraction then takes place at their boundaries.

Values for μ.—For rough calculations it is sufficient to note that the refractive index of water is 4/3 and that of ordinary plate-glass 3/2. For more accurate values see p. 164 and Table I, p. 281. As a result of ideas which have now been abandoned, the custom has arisen of calling the medium with the greater refractive index the optically *denser* medium. For the relationship between the refractive index and the constitution of the substance and the variation of the refractive index with the wave-length of the light see Chap. VII, § 5, p. 163. For X-rays the refractive indices of solids are slightly less than unity (the difference being of the order of a millionth; p. 165).

The Deviation.—The angle through which a light ray is turned as a result of refraction is called the **angle of deviation** or simply the **deviation.** If the angle of incidence is i and the angle of refraction r the deviation δ is equal to $i - r$.

The deviation is greater the greater the angle of incidence.

Proof.—Since $\mu \sin i = \mu' \sin r$, we have $\sin i - \sin r = \{(\mu' - \mu)/\mu\} \sin r$. Hence

$$2 \sin\frac{i-r}{2} \cos\frac{i+r}{2} = \left(\frac{\mu' - \mu}{\mu}\right) \sin r,$$

or, since $i - r = \delta$,

$$2 \sin\frac{\delta}{2} = \frac{\mu' - \mu}{\mu} \cdot \frac{\sin r}{\cos (i + r)/2}.$$

As the angle i increases, r increases also, so that the numerator increases, while the denominator $\cos (i + r)/2$ decreases as i and r increase. Hence the whole expression on the right-hand side, and hence $\sin \delta/2$ and δ itself, increases with the angle of incidence.

The Refraction Law of Möbius.*—Let a ray of light meet the boundary of

* A. F. Möbius (1790–1868), Professor of Mathematics at Leipzig.

two media with refractive indices μ and μ' at A (fig. 29). Then $\mu \sin i = \mu' \sin i'$. Let any straight line cut the normal to the point of incidence at M, the refracted ray at P′, and the incident ray produced at P. From the triangles MPA and MP′A we then have

$$\sin i = \frac{MP}{PA} \sin M \text{ and } \sin i' = \frac{MP'}{P'A} \sin M.$$

Using the equation above we obtain

$$\frac{AP}{AP'} : \frac{MP}{MP'} = \frac{\mu}{\mu'}.$$

If we call $\frac{AP}{AP'} : \frac{MP}{MP'}$ a double ratio, we have the following statement in words:

The double ratio of the intercepts made on the ray by a straight line and by the normal at the point of incidence and the intercepts made on the line by the ray and the normal at the point of incidence, is equal to the ratio of the refractive indices.

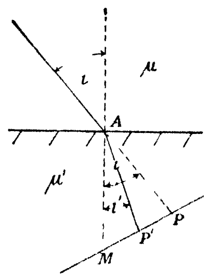


Fig. 29.—To illustrate the proof of Mobius' law of refraction

Construction of the Refracted Ray.—In order to construct the refracted ray corresponding to a given incident ray, we proceed as follows. In fig. 30 OO is the plane boundary surface of the refracting substance. A ray of light LA falls on the boundary at A. We draw the normal at A and describe two circles about A such that the ratio of their radii is equal to the refractive index. In our figure the radii of the two circles are in the ratio of 3 : 2 (p. 54). The incident ray cuts the smaller circle at B. We draw the perpendicular from B to OO, cutting OO at C and the greater circle at D. If we join DA and produce it beyond A, the portion AG represents the path of the refracted ray. For

$$\angle SAL = \angle ABC = i \text{ (the angle of incidence)},$$

and $$\angle S_1AG = \angle CDA = r \text{ (the angle of refraction)}.$$

From the figure we have

$$\sin i : \sin r = \sin CBA : \sin CDA = \frac{AC}{AB} : \frac{AC}{AD} = AD : AB = {}_1\mu_2.$$

If the ray passes from glass to air, the refracted ray is obtained by starting with the straight line GA, finding its intersection (D) with the larger circle, dropping the perpendicular DC, which cuts the smaller circle at B, and joining AB.

Total Reflection.—From fig. 30 we see that the last construction is impossible unless the perpendicular from D on OO cuts the smaller circle. The extreme case accordingly occurs where the perpendicular *touches* the small circle. This limiting case is illustrated separately in fig. 31. In this case the refracted ray travels along the surface of the glass. The *critical angle* γ determining the limiting case is given by the relationship

$$\sin \gamma = \sin ADE = \frac{AE}{AD} = \frac{1}{{}_1\mu_2}.$$

As $\mu = 3/2$ for glass relative to air, the critical angle is given by

$$\sin\gamma = 0{\cdot}67, \text{ i.e. } \gamma = 42^\circ.$$

If the angle of incidence in the glass is greater than this there is no angle of refraction in the air corresponding to it, that is, the ray

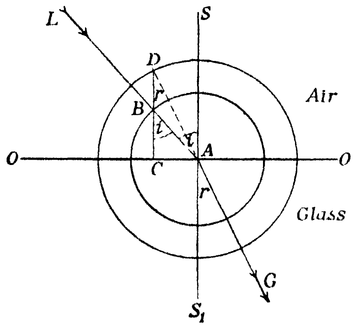


Fig. 30.—Construction for the refracted ray

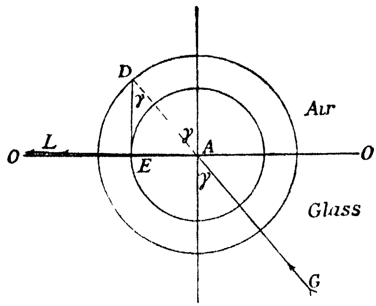


Fig. 31.—Refracted ray grazing the boundary surface

of light cannot leave the glass. When a ray of light is incident on the boundary of two refracting media both refraction and reflection usually take place. If, however, the angle of incidence exceeds the critical angle, the intensity of the refracted ray drops to zero, so that the reflected part of the ray possesses the whole intensity of the incident light. For this reason the phenomenon is called **total reflection.**

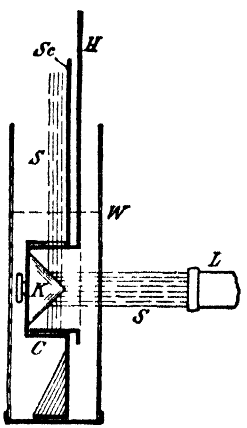


Fig. 32.—Grimsehl's refraction apparatus

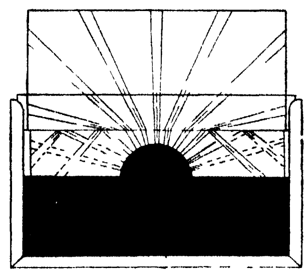


Fig. 33.—Front view of part of the apparatus in fig. 32

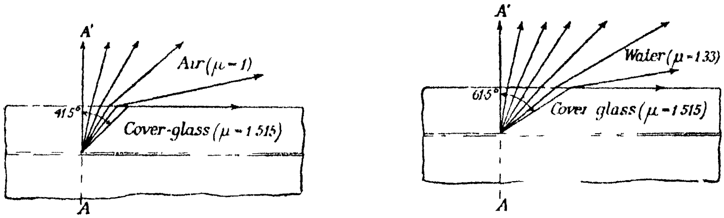
Total reflection invariably occurs when a ray of light from an optically denser medium is incident on the boundary surface of an optically less dense medium at an angle exceeding the critical angle of total reflection determined by the equation $\sin\gamma = 1/{}_1\mu_2$.

If we look obliquely upwards at the horizontal surface of water in a glass, the occurrence of total reflection is signalized by the peculiarly intense reflection, which is as bright as if the light were being reflected by a polished silver plate.

The behaviour of a ray of light which passes from an optically denser medium through the boundary surface of an optically less dense medium may be demon-

strated by the apparatus shown in vertical section in fig. 32. Fig. 33 shows the phenomena observed. A glass tank W filled with water contains a sheet of metal Sc, which has its front painted white and is provided with a cylindrical side chamber of metal, C. To the bottom of the latter there is screwed a polished metal cone K whose axial section is a right-angled isosceles triangle. The cone acts as a mirror. A pencil of parallel rays S emitted by an arc lamp L falls on the vertex and curved surface of the cone, and as a result of reflection is transformed into a pencil of rays starting from the axis of the cone. The rays of light now emerge in all directions from a number of openings in the top of the cylindrical side-chamber and fall on the white screen. The central ray, which meets the surface of the water at right angles, passes through the surface without refraction, while the other rays are subdivided into a reflected ray and a refracted ray. (For four of the pencils in fig. 33 the reflected part has not been drawn.) The outermost rays, which meet the boundary surface at an angle which is greater than the critical angle of total reflection, are totally reflected.

The fact that refraction out of the optically denser medium can only take place for a limited range of angles of incidence on the boundary



Figs. 34, 35.—Pencils of rays emerging from an optically denser medium

may readily be deduced from the fundamental equation $\mu \sin i = \mu' \sin i'$. If $\mu' > \mu$, the greatest value of $\mu' \sin i'$ is μ, as $\sin i$ cannot exceed unity. In this limiting case $\sin i' = \mu/\mu'$ and $i = 90°$. The ray inclined at the angle i' in the optically denser medium therefore grazes the boundary of the two media after refraction. The value of i' for which this takes place is the *critical angle*. If i' exceeds the value corresponding to this case, the fundamental equation can no longer be satisfied and it loses any physical meaning; that is, there is no refraction. The ray is then totally reflected.

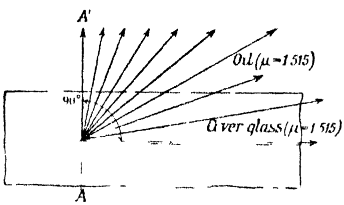


Fig. 36.—A pencil of rays passing through two media with the same refractive index

Aperture of a Pencil of Rays emerging from an Optically Denser Medium.—In connexion with a large number of optical problems, such as the brightness of images, the resolving power, and the depth of focus of the microscope, the following question is of importance: what is the maximum aperture of a pencil of rays which can pass from a luminous point in medium 1 to medium 2 across a plane boundary? Figs. 34–36 show the paths of the rays for some cases which are important in practice (the value of μ for the plate is 1·515). If the upper medium is air, a pencil of rays whose outermost ray makes an angle of 41° 30′ with the normal at the point of incidence can emerge from the lower medium;

if the upper medium is water, the angle is 61° 30′, and if the medium is oil with the same refractive index as the plate, it is 90°. In the figures only the right-hand half of the pencil is shown: the complete conical pencil of rays is to be imagined as formed by rotation about the line AA′. The maximum numerical apertures, namely, those of the outermost rays emerging from the lower medium, are $1{\cdot}515 \,.\, \sin 41° \, 30' = 1$, $1{\cdot}515 \,.\, \sin 61° \, 30' = 1{\cdot}33$, and $1{\cdot}515 \,.\, \sin 90° = 1{\cdot}515$ respectively, that is, equal to the refractive indices of the optically less dense media into which the rays emerge.

The Totally-reflecting Prism.—Total reflection is frequently utilized in order to bend a light ray through a right angle. This is done by means of a right-angled isosceles prism (fig. 37). The light enters the prism at right angles to a side face and is totally reflected at the hypotenuse, as the angle of total reflection for glass relative to air is less than 45°. The light emerges at right angles to the other side face. That is, the totally-reflecting prism is a substitute for reflection at a plane mirror at an angle of 45°, compared with which it possesses the advantage that clear and very bright images are obtained owing to the avoidance of repeated reflection. By reflection at the side faces (indicated by dotted lines in fig. 22, p. 134) the ray may be bent through 180°.

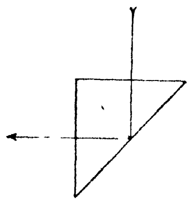


Fig. 37.—Totally-reflecting prism

Path of the Energy in Total Reflection. For NEWTON the existence of total reflection was a weighty reason against the wave theory of light, as he quite correctly showed that on this theory it should be possible to detect a wave motion even beyond the boundary of the optically denser medium. As we now know, the explanation of its non-appearance lies in the interference of the waves which pass into the optically less dense medium. The energy flow in the optically less dense medium at the boundary is accordingly purely tangential (on the average), i.e. along the boundary surface. Nevertheless, it extends into the optically less dense medium in a layer one wave-length thick. That is, the wave motion breaks down into a sort of "surf", but in such a way that the energy passes back entirely into the optically denser medium near the boundary.

According to BEREK the existence of wave motion in the optically less dense medium may be demonstrated very beautifully by covering the boundary with a layer of liquid containing small suspended particles. If we let a pencil of light fall on the boundary from below at an angle exceeding the critical angle and examine the boundary layer with a microscope, we see that the particles close to it diffract the light and are visible as glittering points on a dark ground in virtue of their Brownian movement.

For the relationship between the intensities of the refracted and reflected light and the angle of incidence, see p. 12, fig. 228.

6. Applications of the Law of Refraction to Solids with Plane Boundaries.

Plate bounded by Parallel Planes.—The shading in fig. 38 represents a section through a transparent plate (say of glass) bounded by two parallel planes. A ray of light LA travelling through the air meets the upper surface of the plate at the angle i and is refracted towards the normal at the point of incidence, the angle of refraction in the glass being r. The ray of light then travels on in a straight line and meets the lower surface of the plate at B; the angle of incidence (in the glass) is equal to r, owing to the two surfaces being parallel. On emerging into the air the ray is refracted away from the normal, the angle of refraction (in the air) being equal to i, the original angle of incidence.

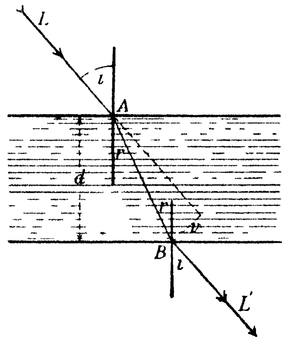


Fig 38.—Parallel displacement due to a plate bounded by parallel planes

This follows immediately from the law of refraction, as owing to the surfaces being parallel $\mu \sin i$ is an invariant for the ray (p 53). That is, the ray of light is merely displaced to one side.

The displacement (v) is greater the greater the thickness of the glass plate (d) and the greater the angle of incidence (i). The construction indicated by the dotted lines in the figure enables us to show that

$$v = d \sin(i - r)/\cos r.$$

The effect of a flat plate on a ray of light may be very well observed by leading a pencil of light rays obliquely upwards (fig. 39, Plate III) through a glass trough full of water and having a sheet of plate-glass for its bottom. If the pencil of light rays is so arranged that part travels through the water and the remainder outside, the magnitude of the parallel displacement is also apparent.

Prisms. *Definitions.*—An optical prism (fig. 40) is a transparent body, two of whose boundaries are planes inclined at an angle to one another. The nature of the remaining parts of the boundary is of no consequence in connexion with the action of the optical prism. The angle W included by the two boundary planes is called the **refracting angle** or simply the **angle** of the prism, and the edge AB along which the two planes meet is called the **refracting edge.** To find the refracting angle and the refracting edge of a prism, it may be necessary to imagine the refracting planes produced. A plane meeting the prism at right angles to the refracting edge is called a **principal section.**

The principal section meets the refracting planes in straight lines whose intersection is that of the principal section with the refracting edge. A perpendicular to the refracting plane at any point of these straight lines lies wholly

within the principal section. A ray of light incident in the principal section on the refracting plane remains in the principal section after refraction (fig. 41).

Path of the Rays in the Principal Section.—In what follows we consider only the path of the rays in the principal section shown in fig. 41. When a ray of light passes from air into a glass prism the ray is bent towards the normal. It then travels on in a straight line in the glass until it reaches the second refracting plane; here it is again refracted, away from the normal. As a result of this repeated refraction the ray is subject to a change of direction or **deviation.**

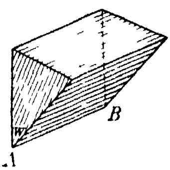


Fig. 40.—A prism

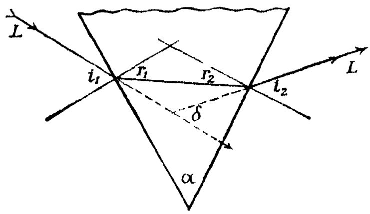


Fig. 41.—Path of a ray through a prism

The deviation δ is the angle which the incident light ray produced makes with the refracted ray produced backwards. If, as is almost always the case, the refractive index of the substance of which the prism is made exceeds that of the surroundings, the ray is bent away from the edge of the prism.

Using the notation of fig. 41, we have

$$\delta = i_1 - r_1 + i_2 - r_2 \text{ and } \alpha = r_1 + r_2,$$

whence $$\delta = i_1 + i_2 - \alpha.$$

The magnitude of the deviation δ therefore depends on the angles i_1 and i_2 and the refracting angle of the prism (α).

Minimum Deviation.—*The deviation caused by the prism is least when $i_1 = i_2$,* or, in other words, *minimum deviation occurs when the path of the ray is symmetrical.*

Proof.—(1) We assume that $i_1 > i_2$, i.e. that $i_1 - i_2 > 0$. Adding the two equations $\sin i_1 = \mu \sin r_1$ and $\sin i_2 = \mu \sin r_2$, we have

$$\sin i_1 + \sin i_2 = \mu (\sin r_1 + \sin r_2).$$

Hence, by transformation,

$$2 \sin \tfrac{1}{2}(i_1 + i_2) \cos \tfrac{1}{2}(i_1 - i_2) = 2 \mu \sin \tfrac{1}{2}(r_1 + r_2) \cos \tfrac{1}{2}(r_1 - r_2),$$

and, as $$r_1 + r_2 = \alpha,$$

$$\sin \frac{i_1 + i_2}{2} = \mu \sin \frac{\alpha}{2} \frac{\cos (r_1 - r_2)/2}{\cos (i_1 - i_2)/2}.$$

But if $$i_1 > i_2,\ i_1 - r_1 > i_2 - r_2 \text{ (p. 54)},$$

so that $$\tfrac{1}{2}(i_1 - i_2) > \tfrac{1}{2}(r_1 - r_2).$$

Since the cosine of an angle diminishes as the angle increases, we have

$$\cos\tfrac{1}{2}(r_1 - r_2) \geqslant \cos\tfrac{1}{2}(i_1 - i_2),$$

i.e.
$$\frac{\cos(r_1 - r_2)/2}{\cos(i_1 - i_2)/2} \geqslant 1.$$

The upper sign applies when $i_1 > i_2$ and the sign of equality when $i_1 = i_2$. Hence the smallest possible value of the above quotient occurs when $i_1 = i_2$. In this case we may put $i_1 = i_2 = i$; then the expression for $\sin\frac{1}{2}(i_1 + i_2)$, which determines δ (see above), also reaches its minimum, namely, $\sin i = \mu \sin \alpha/2$.

(2) If we assume that $i_1 < i_2$, we have $i_1 - i_2 \lesssim 0$, i.e. $r_1 - r_2 \lesssim 0$. The reasoning is the same as in (1), except that $\frac{1}{2}(i_1 - i_2) \lesssim \frac{1}{2}(r_1 - r_2)$. Both values, however, are negative; as regards absolute value, therefore, the difference of the angles of incidence exceeds that of the angles of refraction. The quotient $\cos\frac{1}{2}(r_1 - r_2)/\cos\frac{1}{2}(i_1 - i_2)$ involves the absolute values of the angles only, as the cosine of a negative angle is equal to that of the positive angle of the same magnitude. Hence in this case also we have

$$\frac{\cos(r_1 - r_2)/2}{\cos(i_1 - i_2)/2} \geqslant 1,$$

whence it again follows that the deviation is a minimum for $i_1 = i_2$.

The occurrence of minimum deviation when the path of the light ray is symmetrical may be demonstrated by experiment in the following way.

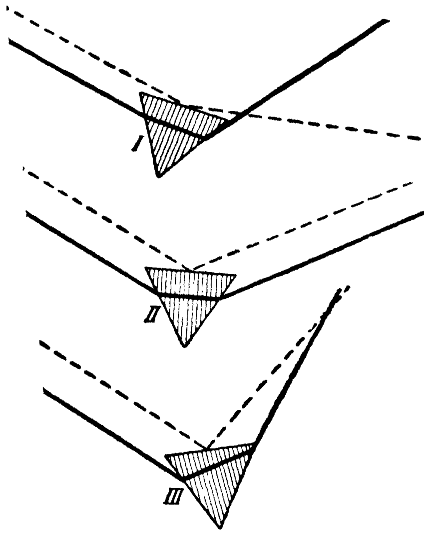


Fig. 42.—Minimum deviation occurs when the path of the ray is symmetrical

We use a prism whose principal section is an equilateral triangle and whose third face is also polished (fig. 42). If a parallel pencil of light is allowed to fall on the prism, part of the pencil is reflected from the third face (as shown by dotted lines in the figure), while another part is deviated by refraction at the two other faces (as shown by continuous lines in the figure). If the prism is now slowly rotated from position I to position II and position III, we find that the refracted part of the pencil is least deviated at the instant when it emerges parallel to the reflected part of the pencil; now this position is characterized by symmetry of the path of the ray.

Determination of the Refractive Index from the Minimum Deviation.—It is possible to make very accurate observations of the minimum deviation. We measure the minimum deviation (δ) and the refracting angle of the prism (α); the reflecting goniometer (fig. 3, Plate III) may be used for both purposes. By p. 61 we have $i = \frac{1}{2}(\delta + \alpha)$ and $r = \frac{1}{2}\alpha$ for the case where the path of the ray is symmetrical, so that

$\sin i = \sin \frac{1}{2}(\delta + \alpha)$ and $\sin r = \sin \frac{1}{2}\alpha$. Dividing and using the equation $\sin i/\sin r = \mu$, we obtain

$$\mu = \frac{\sin \frac{1}{2}(\delta + \alpha)}{\sin \frac{1}{2}\alpha}.$$

for the refractive index.

Deviation by Prisms with a very small Refracting Angle.—If the refracting angle of a prism is very small, δ will also be very small, by the above equation. Hence we may replace the sine by the angle itself (in radians). Then

$$\mu = \tfrac{1}{2}(\delta + \alpha)/\tfrac{1}{2}\alpha,$$

whence we have

$$\delta = (\mu - 1)\alpha\,.$$

If the prism is embedded in a medium of refractive index μ_1 instead of in air, we have to replace μ by the ratio μ/μ_1, by p. 53, and the corresponding formula is

$$\delta = \frac{\mu - \mu_1}{\mu_1}\alpha.$$

In view of the method by which they have been derived these equations are true only when the path of the light is symmetrical, that is, only for small angles of incidence, since $i = (\delta + \alpha)/2$. If the angle of incidence is arbitrary and we bear in mind that $|i_1| \approx |i_2| \approx i$, $|r_1| \approx |r_2| \approx r$; further, that δ, α, and $\delta + \alpha$ are small and that (as is shown by a figure like fig. 41 drawn for a small value of α) i_2 and r_2 have the opposite signs to i_1 and r_1 for fairly large angles of incidence, we obtain by differentiation of $\sin i = \mu \sin r$ the expression

$$\Delta i = i_1 + i_2 = \mu \frac{\cos r}{\cos i}\Delta r = \mu \frac{\cos r}{\cos i}\alpha,$$

whence

$$\delta = \left(\mu \frac{\cos r}{\cos i} - 1\right)\alpha \text{ or } \delta = \frac{\mu \dfrac{\cos r}{\cos i} - \mu_1}{\mu_1}\alpha.$$

When $\cos r/\cos i \approx 1$, therefore, the deviation is independent of the angle of incidence.

CHAPTER IV

Geometrical Optics: Lenses

1. Refraction by Convergent (Convex) Lenses.

Lenses.—A *spherical lens* (fig. 1) is a body consisting of some transparent material which is bounded by two spherical surfaces. If the lens is thicker in the centre than at the edge it is called a convex

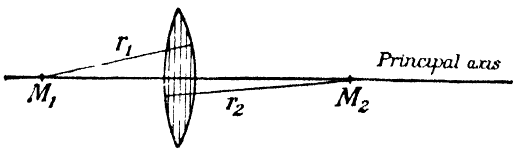


Fig. 1.—Biconvex lens

lens. M_1 and M_2, the centres of the spheres of which the surfaces of the lens form part, are called the **centres of curvature,** and r_1 and r_2, the radii of the spheres, are called the **radii of curvature.** The line joining the centres of curvature is called the **principal axis** of the lens.

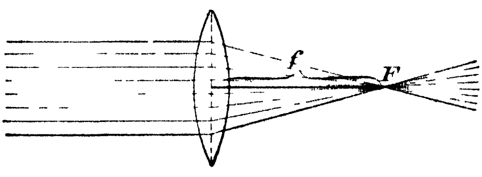


Fig. 2.—Focus of a lens

Observed Facts.—Observation shows that a convex lens renders a parallel pencil of rays convergent (fig. 2). The point where the rays meet is called the **focus** (F). The fact that even divergent rays may be made to converge by means of the convex lens follows from the fact that the convex lens can produce real images of objects situated at a finite distance.

Mode of Action of a Lens.—In order to see how a convex lens acts, we set up the apparatus shown in fig. 3, Plate IV.

Three dishes full of water are placed one on top of the other; the centre one represents a plate bounded by parallel planes, the two others being prism-shaped.

If three parallel pencils of rays are allowed to fall on the dishes, the centre pencil passes through without change in direction, while the upper is deviated downwards and the lower is deviated upwards. The three pencils of rays meet at some distance behind the glass dishes.

If instead of the three prism-shaped dishes we set up a large number of small prisms, of which the outer ones have a larger refracting angle than the inner ones (fig. 4), each prism deviates the incident rays away from its refracting edge, and in fact the outer ones do so to a greater extent than the inner ones. All the rays may be made to meet at one point behind the apparatus. By increasing the number of prisms, we obtain a glass body with curved convex surfaces on either side. If the boundary surfaces are portions of a sphere, the body is a spherical convex lens.

Formation of an Image by a Spherical Refracting Surface. Suppose that two media with refractive indices μ and μ' are separated by a spherical surface (fig. 5). A light ray is incident on the refracting

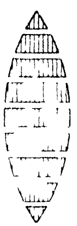


Fig. 4.—Lens regarded as made up of prisms

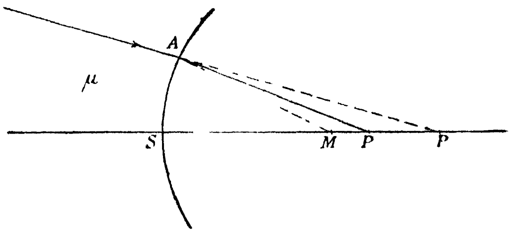


Fig. 5.—To illustrate refraction at a spherical surface

surface at A in the direction AP; let it be refracted in the direction AP′. S is the pole of the refracting surface, M its centre of curvature, and SM = AM = r its radius. Then SM is called the **principal axis** of the spherical refracting surface. The intercepts SP = u and SP′ = v made by the incident and refracted rays on the principal axis and also the radius of curvature SM are to be measured positive from S to the right (in the direction of the light; see p. 40).

We now apply Möbius' theorem (p. 55) to fig. 5. We then have

$$\frac{\mathrm{AP}}{\mathrm{AP'}} : \frac{\mathrm{MP}}{\mathrm{MP'}} = \frac{\mu}{\mu'}.$$

Limitation to Rays near the Axis.—Suppose that a pencil of rays all making a very small angle APS with the principal axis SM falls on the refracting surface, i.e. let A be very close to the pole S (p. 40). Then we may replace AP by SP ($= u$) and AP′ by SP′ ($= v$) in the above equation. If we put MP = SP − SM = $u - r$ and MP′ = SP′ − SM = $v - r$, we have

$$\frac{u}{v} \cdot \frac{v - r}{u - r} = \frac{\mu}{\mu'}.$$

This may be transformed into the **fundamental equation for refraction at a spherical surface,**

$$\mu\left(\frac{1}{r}-\frac{1}{u}\right)=\mu'\left(\frac{1}{r}-\frac{1}{v}\right).$$

This expression is an invariant (the so-called invariant of the surface for axial rays).

The equation may also be written in the form

$$\frac{\mu'}{v}=\frac{\mu}{u}+\frac{\mu'-\mu}{r}.$$

Refractive Power.—The expression $(\mu'-\mu)/r$ is called the *refractive power* of the surface. Its unit is called the **dioptre.** Here r is to

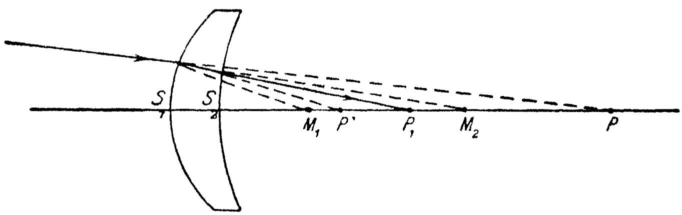


Fig. 6.—Path of the rays for a convexo-concave lens

S_1 and S_2 are the points where the principal axis meets the lens, M_1 and M_2 the centres of curvature of its surfaces; P is the (virtual) object, P′ the image formed by the front surface and P_1 the image formed by the lens as a whole.

be measured in metres; a surface accordingly has a power of one dioptre if the difference of the refractive indices on the two sides of the surface is equal to the radius of curvature expressed in metres.

Formation of an Image.—The above equation connecting u and v is independent of the inclinations of the incident and refractive rays to the axis. All the incident rays directed towards P which satisfy the limitation stated at the beginning of the discussion will accordingly pass through P′. P′ is therefore the (real) image of the (virtual) object P. The fundamental equation stated above accordingly gives the relationship between the positions of the object and the image.

If the first medium is air, $\mu=1$, and the equation takes the form

$$\frac{\mu'}{v}=\frac{1}{u}+\frac{\mu'-1}{r}; \qquad \ldots\ldots\ldots \qquad (a)$$

if, on the other hand, the second medium is air, $\mu'=1$, and the equation takes the form

$$\frac{1}{v}=\frac{\mu}{u}+\frac{1-\mu}{r}. \qquad \ldots\ldots\ldots \qquad (b)$$

Formation of an Image by a Lens (fig. 6).—Let the principal axis

of the lens meet the spherical surfaces at S_1 and S_2. Let there be air to the left of S_1 and to the right of S_2 and let the refractive index of the lens be μ. P is the virtual object relative to the front surface, i.e. relative to the lens; P′ is the image which would be produced by this surface alone if the medium behind it were of refractive index μ. We then have the fundamental equation

$$\frac{\mu}{S_1P'} = \frac{1}{S_1P} + \frac{\mu - 1}{S_1M_1}.$$

For the second surface P′ is the object; P_1 is the image of it formed by the second surface and hence the image of P formed by the lens as a whole. We have

$$\frac{1}{S_2P_1} = \frac{\mu}{S_2P'} + \frac{1-\mu}{S_2M_2}.$$

Further, let S_1S_2, the *thickness of the lens*, be d, and let $S_1M_1 = r_1$, $S_2M_2 = r_2$, $S_1P = u$, $S_2P_1 = v$. Then addition of the two equations gives

$$\frac{1}{v} + \frac{\mu}{S_2P' + d} = \frac{1}{u} + \frac{\mu}{S_2P'} + (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right).$$

Limitation to Infinitely Thin (Ideal) Lenses.—If d is vanishingly small compared with S_2P', we obtain the **object-image equation for a thin lens,**

$$\frac{1}{v} = \frac{1}{u} + \left(\frac{1}{r_1} - \frac{1}{r_2}\right)(\mu - 1).$$

In this equation, by the method of derivation and the sign convention adopted (see above) u is *negative* for a *real* object, *positive* for a *virtual* object. On the other hand, v is *positive* for a *real* image, *negative* for a *virtual* image.

For the so-called *bi-convex lens* (fig. 1), r_2 is negative, as the centre of curvature lies to the left of the lens. A lens for which r_1 is infinite and the centre of curvature of the second surface is to the left of the lens, is called a *plano-convex lens.* A lens like that shown in fig. 6, in which the centres of curvature of both surfaces lie to the right of the first surface and $r_2 > r_1$, is called a *convexo-concave* * *lens* or *meniscus.*†

Foci of a Lens.—The point where parallel rays falling on the lens are reunited is called the **focus.** Its distance from the lens is called the **focal length** of the lens; we shall denote it by f_1. When u is infinite,

* [The surfaces of the lens are referred to in the order in which the light reaches them.]

† Gr., *mēniskos,* little moon, a diminutive of *mēne,* the moon; i.e. here meaning with a cross-section shaped like a new moon.

$v = f_1$. Substituting these values in the equation above, we have

$$\frac{1}{f_1} = 0 + (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

or

$$f_1 = \frac{r_1 r_2}{(\mu - 1)(r_2 - r_1)}.$$

If v is infinite, we obtain a value for u which we shall denote by f, $f = -f_1$. That is, a lens (in contradistinction to a concave mirror) possesses two distinct foci; one, the *first principal focus*, lies to the left of the lens, while the other, the *second principal focus*, lies to the right of the lens (at a distance equal to the focal length f_1). The two focal lengths are equal and of opposite sign.

If we substitute the focal length f_1 in the object-image equation, it takes the form

$$\frac{1}{v} = \frac{1}{u} + \frac{1}{f_1},$$

which is the same as that which we obtained for the concave mirror.

The Convergence Equation.—Accordingly $D = 1/f_1$ is the refractive power of the lens. Just as in the case of the concave mirror (p. 42) we may call $U (= 1/u)$ and $V (= 1/v)$ the *convergence of the object* and the *convergence of the image*, both being measured in dioptres. We then have $V = U + D$, that is, the convergence of the image is the sum of the convergence of the object and the power of the lens.

In the particular case where the two radii of curvature of the lens are equal and opposite (we then put $r = r_1 = -r_2$) and the lens is composed of glass of refractive index $\mu = 3/2$, we have $f_1 = \frac{1}{3/2 - 1} \cdot \frac{r^2}{2r} = r$ (Kepler). Hence a lens of this glass with radii of curvature 1 metre has a power of 1 dioptre.

Newton's Form of the Equation.—If we replace the distances u and v in the equation $1/v = 1/u + 1/f_1$ by the expressions $u = g - f_1$ and $v = b + f_1$, where g, b, are the *focal distances* (measured from the appropriate foci (fig. 8)), the lens formula becomes

$$bg = -f^2 = -f_1^2 = ff_1,$$

which is Newton's form of the object-image relationship.

These equations may be interpreted and illustrated graphically, just as was done on p. 43 for the concave mirror.

Subsidiary Axes.—Our deduction of the equation for a thin lens is subject to the proviso that the dimensions of the lens must be small in comparison with the distances measured along the principal axis, so that any rays not in the direction of the principal axis are inclined to it at a small angle only.

For prisms with a small angle and small angles of incidence, the deviation of a ray is independent of the angle of incidence (p. 62).

Hence if we allow a pencil of parallel rays inclined at a small angle to the principal axis to fall on a lens of small curvature, these parallel

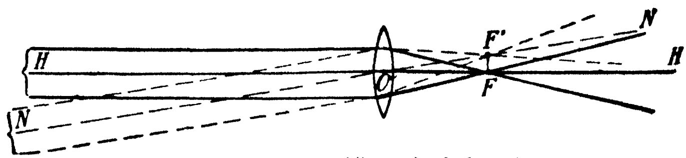


Fig. 7.—Principal axis and subsidiary axis of a lens; focal plane

rays are also reunited in one point behind the convex lens. A straight line inclined to the principal axis and passing through O, the centre of the lens, may be called a **subsidiary axis.** (In fig. 7 HH is the principal axis, NN a subsidiary axis.) The point where the rays parallel to the subsidiary axis are reunited is the point F′, which must lie on the subsidiary axis itself, as (apart from the parallel displacement, which owing to the infinitesimal thickness of the lens is infinitesimal) the direction of the axis is not changed in traversing the lens. F′ lies as

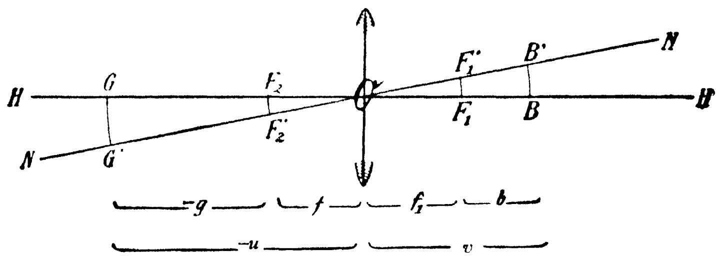


Fig. 8.—Object plane and image plane

far behind the lens on the subsidiary axis as F does on the principal axis; hence we obtain the position of F′ by imagining the principal axis, together with all the points lying on it, rotated about O, the centre of the lens. FF′ is a small circular arc, which in view of our proviso may be imagined as replaced by a straight line perpendicular to the principal axis. Any subsidiary axis behaves in the same way as that drawn in the figure; that is, for every subsidiary axis there is a focus (one on each side of the lens), to which all rays parallel to the subsidiary axis converge after passing through the lens. All these foci lie on a plane through F at right angles to the principal axis. This plane is called a **focal plane** of the lens.

In the following figures the dimensions of breadth have been considerably exaggerated in comparison with the dimensions of length, as otherwise the figures could not be made clear in the space available. As a result the angles at which the subsidiary axes as drawn are inclined to the principal axis also appear too large. Hence the errors which according to our proviso are to be neglected are more marked than when thin lenses are actually in use. Similarly, the focal

planes, which we shall always indicate by dotted lines, are shown to far too great an extent. In the actual use of lenses it is only the central part of the focal planes that is of any importance. In future, moreover, we shall always represent a lens as in fig. 8, by a line perpendicular to the principal axis, with shaded arcs at the top and bottom to make the nature of the lens clear.

Just as there are two foci F_1 and F_2 on the principal axis HH (fig. 8) and every point source of light or point object G has its image-point B, at which the rays of light starting from G converge after passing through the lens, so every subsidiary axis has corresponding foci, object-points, and image-points. These are obtained by rotating the principal axis about O, the centre of the lens, until it occupies the position of the subsidiary axis. The arcs GG′ and BB′ formed in this rotation may again be regarded as straight lines at right angles to the principal axis; we accordingly call the plane at right angles to the principal axis and passing through G the **object plane** and that

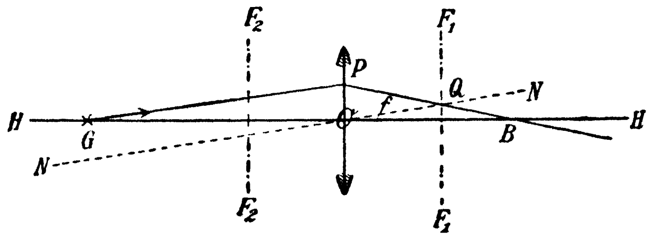


Fig. 9 —Construction for the image formed by a lens

passing through B the **image plane.** The region, starting from the lens, in which the object is situated is called the **object space** and the region in which the image lies the **image space.**

The focal plane on the *right,* which passes through F_1, belongs to the object space on the *left,* and the focal plane on the *left,* which passes through F_2, belongs to the image space on the *right.*

Linear Magnification.- The ratio $m = BB'/GG'$ (fig. 8) is called the linear* lateral magnification. Now $BB'/GG' = OB/OG = v/u$. Hence the linear magnification m is equal to v/u, i.e. equal to the ratio of the distances of the image and the object from the lens.

Construction for the Image formed by a Lens.—A luminous point G lies on the principal axis HH (fig. 9) of the convex lens O with the focal planes F_1F_1 and F_2F_2. An arbitrary ray starting from G meets the lens at P. We draw the subsidiary axis NN parallel to PG, cutting the focal plane F_1F_1 in Q. As after refraction the ray GP parallel to the subsidiary axis intersects this axis in the focal plane in the image space, the refracted ray must travel from P through Q; let it cut the principal axis at B. Any other ray starting from Q will behave in the same way; B is therefore the real image of G.

* To distinguish it from the magnification of area, which is equal to the square of the linear magnification (p. 96).

Note.—As GP is parallel to OQ, we have GP/OQ = GB/OB. If OG = u, OB = v, and the focal distance OQ = f_1, GO = − OG = $-u$, and we accord-

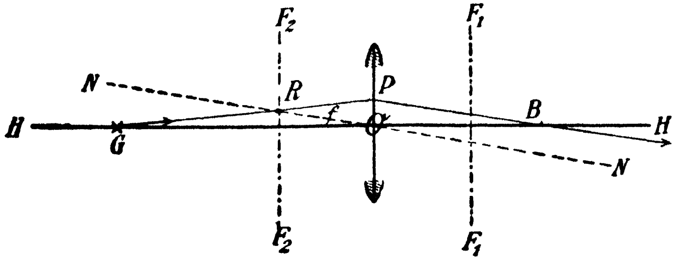


Fig. 10.—Another method of constructing the image

ingly have the equation $-u/f_1 = (-u + v)/v$. By transformation we have $1/v = 1/u + 1/f_1$ (p. 67). That is, the construction is consistent with the object-image formula.

Fig. 10 gives another construction for the image. A ray GP starting from G cuts the focal plane F_2F_2 in the object space at R, so that after traversing the lens it moves parallel to the subsidiary axis NN which passes through R. We accordingly draw the subsidiary axis NN through R and O and through P draw the parallel to NN; this is the refracted ray, which cuts the principal axis HH at B. B is the real image of G.

Note.—As RO is parallel to PB, we have RO/PB = GO/GB; hence, as

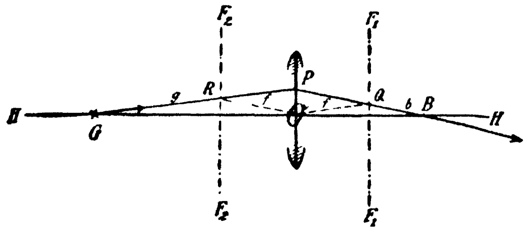


Fig. 11.—Another method of constructing the image

RO = − OR = $-f = f_1$, we have $f_1/v = -u/(-u + v)$. On transformation this again yields the equation $1/v = 1/u + 1/f_1$.

Combining the constructions of figs. 7 and 10, we obtain fig. 11, from which we see that the triangles GRO and OQB are similar. Hence the construction may also be carried out as follows.

Let any ray GP cut F_2F_2 at R. Join RO, draw through O the straight line OQ parallel to GP cutting F_1F_1 at Q, and finally draw QB parallel to RO. Then B is the image of G.

Note.—As the triangles GRO and OQB are similar,

$$\frac{\mathrm{GR}}{\mathrm{RO}} = \frac{\mathrm{OQ}}{\mathrm{QB}}.$$

If in due agreement with the previous notation we denote the distance of the object from the focal plane F_2F_2 by g and the distance of the image from the focal plane F_1F_1 by b, the above statement of proportion becomes

$$-\frac{g}{f_1} = \frac{f_1}{b}$$

or $$gb = -f_1^2,$$

which is Newton's form of the equation for a thin lens.

We see at once from the completed figures that if G coincides with the focus,

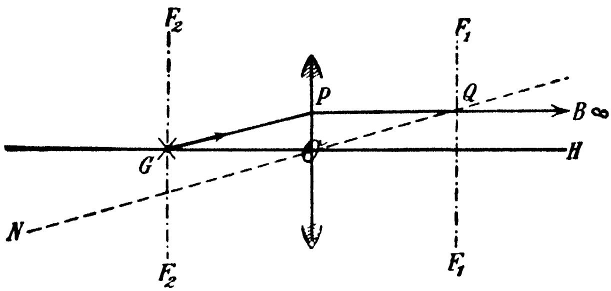


Fig. 12.—Object at the focus

the refracted ray must leave the lens parallel to the principal axis, for in this case GPQO is a parallelogram (fig. 12).

Further, it follows immediately that if G moves to a position between the

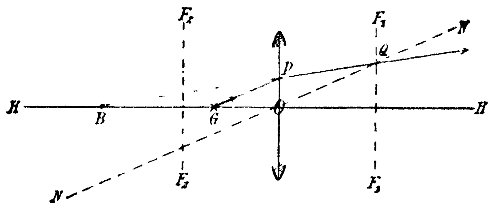


Fig. 13 —Object nearer the lens than the focus

focus and the lens, the refracted ray as it leaves the lens must diverge from the principal axis, for in this case GP < OQ (fig. 13). Now as GP is parallel to OQ, the two straight lines GO and PQ joining their extremities must intersect to the left of G. The point of intersection B must lie on the backward prolongation of the ray PQ leaving the lens; that is, B is a virtual image of G.

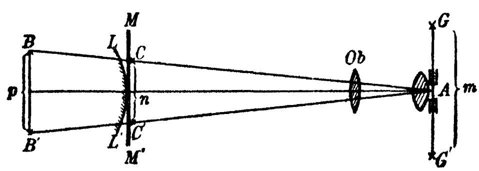


Fig. 14.—Measurement of the radius of curvature of a lens

Optical Method for Measuring the Radius of Curvature of a Convergent Lens.—Two sources of light approximating as nearly as possible to point sources, e.g. two small electric lamps G and G′, are set up on a stand at a definite distance (m) apart (fig. 14). The lens LL′, whose radius of curvature is to be measured, is set up vertically at some distance away (its distance from the stand being x). The stand carrying the lamps is at right angles to the principal axis of the lens.

The front surface of the convergent lens acts as a convex mirror and produces two images B, B′ of the two point sources of light, at a distance of y behind the lens. If x and y could be measured, r, the radius of curvature of the front surface of the lens, could be calculated from the equation for a convex mirror (p. 47), which is

$$\frac{1}{y} = -\frac{1}{x} - \frac{2}{r};$$

y, however, is not accessible to direct measurement. We therefore set up a small scale MM′ just in front of the surface of the lens and observe the apparent magnitude CC′ ($=n$) of the interval BB′ ($=p$) on the scale, possibly with the aid of a telescope (p. 129) set up so that the eyepiece is between G and G′. GG′, the size of the object, BB′, the size of the image, and their distances x and y from the lens are connected by the relationship which holds for all mirrors and lenses (see e.g. p. 69),

$$\frac{\text{Size of object}}{\text{Size of image}} = \frac{\text{Distance of object}}{\text{Distance of image}},$$

so that

$$\frac{\text{GG}'}{\text{BB}'} = \frac{-x}{y}$$

or

$$\frac{m}{p} = -\frac{x}{y}.$$

We do not, however, observe p, the size of the image, directly, but only its projection CC′ (n) on the scale MM′ placed immediately in front of the lens. (If a telescope is used, the centre of projection is the "exit pupil" of the telescope, i.e. the image of the objective produced by the eyepiece, by p. 130). By geometry we have

$$\frac{p}{n} = \frac{(-x + y)}{-x}.$$

Eliminating y and p by using the equations

$$\frac{1}{x} + \frac{1}{y} = \frac{2}{r},$$

$$\frac{m}{p} = -\frac{x}{y},$$

we obtain

$$r = \frac{2nx}{m - 2n};$$

the quantities m, n, and x can be measured.

2. Divergent (Concave) Lenses.

If the centre of a spherical lens (i.e. a lens bounded by spherical surfaces) is thinner than the sides, the lens is called a concave lens. In general it acts as a divergent lens (in this connexion see p. 89).

We shall confine our attention to ideal concave lenses, i.e. to lenses of infinitesimal thickness, and shall deal only with axial rays, i.e. rays inclined at such a small angle to the axis that the angle in radian measure, its sine, and its tangent are interchangeable.

Observation shows (fig. 15) that parallel rays of light (e.g. sunlight) incident on the lens in the direction of the principal axis diverge as

a result of refraction by the lens in such a way that their backward prolongations pass through one point on the principal axis. The rays accordingly appear to diverge from this point, which is called the (virtual) focus of the concave lens; its distance from the lens is called the focal length and is to be reckoned negative (i.e. of the opposite sign to that for the convex lens) as it lies on the same side of the lens as the source of the rays.

The formation of an image by a concave lens takes place entirely according to the laws which we already know, for the process depends on the formation of images by two successive spherical surfaces (p. 64). Hence the equation giving the relationship between the positions of the image and object for the concave lens will follow in exactly the same way as that for the convex lens.

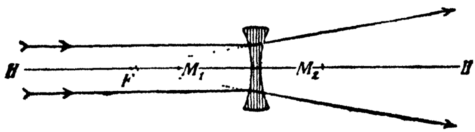


Fig. 15.—Divergent (bi-concave) lens

Whether we are dealing with a convex lens or a concave lens is determined merely by the sign of the focal length. For the convex lens we obtained

$$\frac{1}{f_1} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

(p. 67. If $r_1 > r_2$, i.e. if the first surface of the lens (moving from left to right) has a smaller curvature than the second surface, f_1 is negative; the second focus becomes virtual and the lens is a divergent lens.

If the two centres of curvature lie on the same side of O, i.e. if the curvatures have the same sign, the lens is said to be *concavo-convex* (a meniscus, cf. p. 66); if r_1 is infinite, so that the first surface of the lens is plane, and r_2 is positive, the lens is said to be *plano-concave*; finally, if r_1 is positive and r_2 is negative, so that the centre of curvature of the first surface lies in front of the lens and the centre of curvature of the second surface lies behind it, the lens is said to be *bi-concave* (fig. 15).

In the equation

$$\frac{1}{v} = \frac{1}{u} + \frac{1}{f_1}$$

which is also true for the concave lens, f_1 accordingly has a negative value.

Owing to this negative value of f_1, a negative value of v corresponds to any *real object-point*, as then u has a negative value. Hence v is less than f_1 in absolute value.

In the case of a concave lens, any real object gives rise to a virtual image nearer the lens than the focus.

The construction of the image of a point source G is obvious from fig. 16. Let G be a point on the principal axis HH of the concave lens situated at O, and let F_1F_1 and F_2F_2 be the focal planes of the lens. It is to be noted that the focal plane of a divergent lens corresponding to the rays falling on the lens from a certain side lies on the side from which the rays approach. Let any ray starting from G meet the lens at P. Parallel to GP we draw the subsidiary axis NN cutting the focal plane F_1F_1 at Q. The ray GP parallel to the subsidiary axis NN leaves

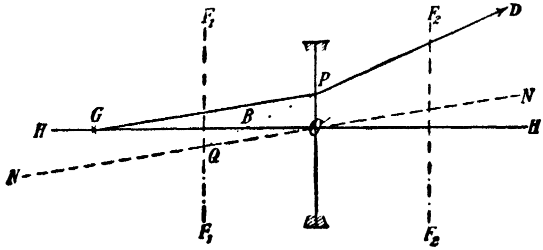


Fig. 16.—Construction for the image formed by a concave lens

the lens in such a direction that it appears to come from Q. We draw QP and produce it beyond P to obtain the subsequent direction of the light ray, PD. B, the point where QP meets the principal axis HH, is the virtual image of G.

3. Image, formed by a Spherical Surface, of a Point on its Axis.

Optical Systems.—In the previous sections we have deduced the path of a ray through a single lens under the assumption that the lens is very thin and hence that the theorems about the refraction of light in a prism of small angle may be used. We have also assumed that the lens is bounded by air on either side. In actual fact, however, lenses are *not* extremely thin; moreover, they often consist of a number of substances with differing refractive indices, a number of lenses are frequently combined to form a single optical system, and, finally, the path of the ray may begin and end in substances of differing refractive index.

Fundamental Equations for a Refracting Spherical Surface.—In order to trace out the paths of the rays through a system of refracting surfaces, we go back to the **fundamental equation for refraction at a spherical surface,** which we deduced on p. 65:

$$\mu\left(\frac{1}{r}-\frac{1}{u}\right)=\mu'\left(\frac{1}{r}-\frac{1}{v}\right),$$

or

$$\frac{\mu'}{v}=\frac{\mu}{u}+\frac{\mu'-\mu}{r}.$$

If r is infinite, $v = u\mu'/\mu$. That is, if we look perpendicularly (or nearly so) at a plane refracting surface, an object lying at a distance u behind the refracting surface appears to be brought nearer, to the distance $u\mu'/\mu$. In practical cases

μ' is usually equal to unity; thus, for example, the bed of a stream or lake appears raised, the ratio of the true depth to the apparent depth being equal to the refractive index of water ($v/u = 1/\mu = \frac{3}{4}$ (p. 52)). This property may be utilized to measure refractive indices. If (e.g. in air) we observe a fine mark on the upper side of a plate bounded by parallel planes, of refractive index μ and thickness u, by means of a microscope, and then lower the microscope through the distance v, until it is sharply focused on a mark on the lower surface of the plate, then $\mu = u/v$.

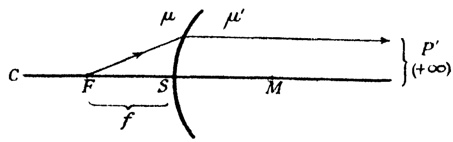


Fig. 17.—The first principal focus

Focal Lengths.—The object-point corresponding to an infinitely distant image ($v = \infty$) is called the focus in the object space or **first principal focus.** In future we shall denote it by F (fig. 17). Its distance from the first surface of the optical system we shall call the focal length in the object space or **first focal length.** This we shall denote by f. If we put $u = f$ and $v = \infty$ in the fundamental equation, we have

$$0 = \frac{\mu}{f} + \frac{\mu' - \mu}{r},$$

or

$$f = -\frac{\mu r}{\mu' - \mu}.$$

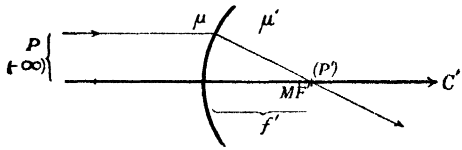


Fig. 18.—The second principal focus

The minus sign indicates that the focus F lies to the left of the refracting surface, in the medium of refractive index μ.

If P (fig. 18) moves to an infinite distance, $u = -\infty$, so that the second term of the fundamental equation vanishes. The position which the image-point P′ then takes up is called the focus in the image space or **second principal focus**; we denote it by F′ and call its distance from S (fig. 18) the focal length in the image space, or **second focal length**; this we shall denote by f'. We obtain this focal length by putting $u = -\infty$ and $v = f'$ in the fundamental equation. We thus obtain

$$\frac{\mu'}{f'} = \frac{\mu' - \mu}{r},$$

so that

$$f' = \frac{\mu' r}{\mu' - \mu}.$$

As this value is positive, the second focus lies in the second medium, which has the refractive index μ'.

If we divide all the terms of the fundamental equation by $(\mu' - \mu)/r$ it becomes

$$\frac{\mu' r}{\mu' - \mu} \cdot \frac{1}{v} = \frac{\mu r}{\mu' - \mu} \cdot \frac{1}{u} + 1.$$

Inserting the values for f and f' in this expression, we obtain

$$\frac{f}{u} + \frac{f'}{v} = 1.$$

If $u < f$, the first term is greater than the right-hand side, and the second term must be negative. It follows that v must be negative, i.e. that the image-point must lie to the left of S, i.e. in the first medium. This is only possible if the rays starting from the object diverge after refraction by the spherical surface. Hence in this case the image is virtual (fig. 16).

The Ratio of the Focal Lengths.—Forming the quotient of the expressions for f and f', we obtain the equation

$$\frac{f}{f'} = -\frac{\mu}{\mu'}.$$

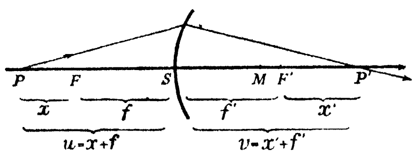


Fig. 19.—To illustrate Newton's form of the fundamental equation for a spherical refracting surface

The absolute values of the two focal lengths of the spherical refracting surface are in the ratio of the refractive indices of the media separated by the spherical surface.

The fundamental equations in the form deduced possess general validity, no matter what the position of the points involved. They also hold for the case where the *centre of curvature of the sphere lies to the left of the refracting surface*, i.e. in the medium with the smaller refractive index. We have then only to insert negative values for r. Here we shall not discuss the positions of image and object in detail, as the argument follows closely on the lines of that given previously. We would merely point out that in this case the first focus lies in the second medium and the second focus in the first medium, and, moreover, that to a real object situated in the first medium there invariably corresponds a virtual image likewise situated in the first medium.

Newton's Form of the Fundamental Equation (p. 67).—The fundamental equation assumes a particularly simple form if we make the focal distances of the object and image the variables. We denote these by x, x' and reckon them positive in the direction in which the light travels. By fig. 19 we then have to put

$$u = x + f \text{ and } v = x' + f',$$

where u, f, and x are negative quantities in fig. 19. We obtain

$$\frac{f}{x+f} + \frac{f'}{x'+f'} = 1,$$

or

$$xx' = ff'.$$

4. Image, formed by a Spherical Surface, of an Object near its Axis.

A Small Arc as Object.—If in fig. 5 (p. 64) we imagine the straight line which passes through M, the centre of the sphere, and on which P and P′ lie, rotated through a small angle, then P and P′ describe small circular arcs, which are shown dotted in fig. 20. As the angles are small, we may replace the arcs by the tangents PQ and P′Q′. Just as the object-point P lying on the axis has an image-point P′ corresponding to it which lies on the axis, the object-point Q near the axis must give rise to an image-point Q′ near the axis. The extended image P′Q′ then corresponds to the extended object PQ.

That is:

The image of a small object PQ *at right angles to the axis* SM *is a small object* P′Q′ *at right angles to the same axis.*

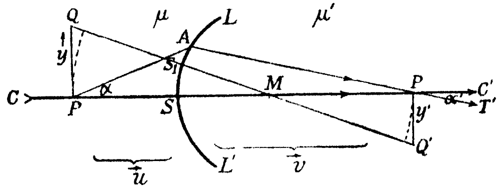


Fig. 20.—Formation of an image of a finite object by a spherical surface

The Helmholtz-Lagrange Theorem.—We shall call the angle at which a ray PA starting from the point P of the axis leaves the axis the **inclination of the ray.** If (fig. 20) we denote the angle SPA by α, the angle C′P′T′ by α', and SA by s, we have

$$\tan\alpha = \frac{\mathrm{SA}}{\mathrm{PS}} = \frac{s}{-u} \text{ and } \tan\alpha' = \frac{\mathrm{SA}}{\mathrm{P'S}} = \frac{s}{-v}$$

or

$$\mathrm{SP} = -\frac{\mathrm{SA}}{\tan\alpha} \text{ and } \mathrm{SP'} = -\frac{\mathrm{SA}}{\tan\alpha'},$$

since SA may be regarded as coinciding with the tangent at the point S (see above) in view of the limitation to axial rays.

As the following considerations apply exclusively to *rays near the axis*, the angle and its tangent are interchangeable. Hence in what follows we shall always use the angle (measured in radians).

By Möbius' theorem (p. 55) we have

$$\frac{\mathrm{SP}}{\mathrm{SP'}} : \frac{\mathrm{MP}}{\mathrm{MP'}} = \mu : \mu';$$

hence

$$\frac{\alpha'}{\alpha} : \frac{\mathrm{MP}}{\mathrm{MP'}} = \frac{\mu}{\mu'}.$$

We call PQ $= y$ the **size of the object** and P′Q′ $= y'$ the **size of the image.** From fig. 20 we see at once that MP/MP′ = PQ/P′Q′ $= y/y'$. Hence the above equation becomes

$$\frac{\alpha'}{\alpha} : \frac{y}{y'} = \frac{\mu}{\mu'},$$

or

$$y\mu\alpha = y'\mu'\alpha'$$

(Helmholtz's equation.*)

* Helmholtz (1856) was the first to recognize the general applicability of this equation to the formation of images by optical systems.

If an image of an object at right angles to the axis is formed by a spherical surface separating two optically differing media, the product of the size of the image, the angle of inclination of the ray measured in radians, and the refractive index is constant. This theorem holds only for *small* angles (axial rays). For larger angles its place is taken by the sine condition (p. 93).

The product $y\mu\alpha$ is an *invariant* for images formed by axial rays.

The Lateral Magnification.—The quotient of the size of the image P′Q′ (y') and the size of the object PQ (y) is called the **lateral magnification.** It is denoted by m: $m = y'/y$.

We write Helmholtz's equation in the form

$$\frac{y'}{y} = \frac{\mu\alpha}{\mu'\alpha'}$$

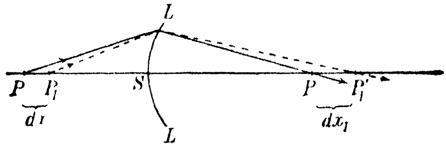


Fig. 21.—To illustrate the elongation of an image

and put $$\alpha = -\frac{s}{u} \text{ and } \alpha' = -\frac{s}{v} \text{ (see above).}$$

Then $$m = \frac{y'}{y} = \frac{\mu v}{\mu' u}.$$

By p. 76 we also have

$$\frac{\mu}{\mu'} = -\frac{f}{f'}$$

and $$v = x' + f', \quad u = x + f.$$

Hence $$m = -\frac{f(x'+f')}{f'(x+f)} = -\frac{fx'+ff'}{f'(x+f)}.$$

By p. 76, we may replace ff' in the numerator by xx'; then

$$m = -\frac{fx'+xx'}{f'(x+f)} = -\frac{x'(f+x)}{f'(x+f)} = -\frac{x'}{f'} = -\frac{f}{x}.$$

The Elongation.—If an object-point P (fig. 21) is moved through a short distance dx along the axis to P_1, the corresponding image-point P′ moves through a distance dx_1, say, to P_1'. The ratio of the displacements along the axis of the image and object respectively may be called the **elongation** (t).

The elongation may be calculated very simply from Newton's form of the fundamental equation, $xx' = ff'$. By differentiating this we obtain

$$x\,dx' + x'\,dx = 0.$$

Hence $$t = \frac{dx'}{dx} = -\frac{x'}{x}.$$

Multiplying above and below by x and replacing the product xx' in the numerator by ff', we have

$$t = -\frac{ff'}{x^2} = -\frac{x'^2}{ff'}.$$

For the case where there are many refracting surfaces (see the following section) and the refractive indices of the first and last media are the same, the first and second focal lengths are equal and opposite. The expression for the elongation then takes the simpler form $t = f^2/x^2$. Comparing this expression with the expression for the lateral magnification $m = -f/x$, we have the following result:

When the media surrounding the object and the image have the same refractive index, the elongation is equal to the square of the lateral magnification.

It follows that the image of an extended object must be distorted, as it is magnified or diminished to a much greater extent along the axis than at right angles to it. Further, we see that it is impossible to obtain a clear image of all points of an extended object on a plane screen. If we photograph an extended object, e.g. a landscape, both foreground and background become sharper the smaller the image is (the principle of the small-sized camera).

The Convergence Ratio or Angular Magnification.—The ratio of the tangents of the inclinations of the rays before and after refraction is called the *convergence ratio* k. We found previously that $\alpha' = -s/v$ and $\alpha = -s/u$. Hence $k = \alpha'/\alpha = u/v = \mu/\mu' m$, or $km = \mu/\mu'$. Replacing the first quotient μ/μ' by $-f/f'$ and m by $-x'/f'$ or $-f/x$, we have

$$k = \frac{x}{f'} = \frac{f}{x'}.$$

The Subjective Magnification. An object y at right angles to the visual axis at a distance u from the eye appears to subtend at the latter a definite natural **visual angle** ϕ, also called the **apparent magnitude**; this angle is determined by the equation $\phi = y/u$. If a system of lenses is interposed between the eye and the object, the apparent magnitude of the object is usually altered. The new visual angle is the angle at which the ray starting from the outermost point of the object is inclined to the axis after traversing the system of lenses. If this angle is ψ, the quotient ψ/ϕ is called the *subjective magnification* (w) of the object due to the system of lenses.

5. Image formed by a Coaxial System of Spherical Surfaces.

Image of a Point on the Axis.—In fig. 22 the spherical surfaces L_1, L_2, L_3 with their centres M_1, M_2, M_3 all lying on the same axis are shown, and also the four media with refractive indices μ_0, μ_1, μ_2,

μ_3, which are separated by the spherical surfaces. Spherical surfaces arranged in this way are said to be **coaxial.** The points where the axis meets the surfaces are denoted by S_1, S_2, and S_3; $S_1S_2 = d_1$, $S_2S_3 = d_2$.

Let P_0 be a luminous point on the axis. The spherical surface L_1 (subject to limitation to axial rays) gives rise to the image-point P_1 lying on the axis. This serves as object for the spherical surface L_2, which in its turn gives rise to the image-point P_2. Again, P_2 acts as object for the surface L_3, which gives rise to the image-point P_3. Thus we may imagine a whole series of spherical surfaces, n in all, with the same axis. The system of coaxial spherical surfaces then gives rise to an image P_n in the last medium of the object P_0 in the first medium.

A pencil of rays starting from a point source (homocentric pencil) is

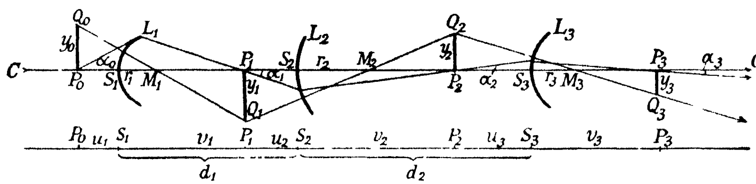


Fig. 22.—Coaxial system of spherical surfaces

transformed into another homocentric pencil by a system of coaxial spherical surfaces separating media of differing refractive index (provided all the rays are axial).

The position of the last image P_n may be calculated if the radii of curvature of the spherical surfaces, the distances between their centres or poles, and the refractive indices of the media are known.

The actual carrying-out of the calculation necessitates writing down the fundamental equation (deduced on p. 65) for each surface and hence evaluating v_n, the distance between the last surface and the final image, using the distances between the poles of the surfaces.

The system of equations would be as follows:

$$\frac{\mu_1}{v_1} = \frac{\mu_0}{u_1} + \frac{\mu_1 - \mu_0}{r_1}$$

$$\frac{\mu_2}{v_2} = \frac{\mu_1}{u_2} + \frac{\mu_2 - \mu_1}{r_2}$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{\mu_n}{v_n} = \frac{\mu_{n-1}}{u_n} + \frac{\mu_n - \mu_{n-1}}{r_n}.$$

Image of an Object at right angles to the Axis. If at P_0 there is an object P_0Q_0 of magnitude y_0 at right angles to the axis, there result the successive images P_1Q_1, P_2Q_2, . . ., P_nQ_n at right angles to the axis and of magnitude $y_1, y_2, \ldots, y_n$. By p. 78 the **lateral magnification** is

$$m = \frac{y_n}{y_0} = \left(\frac{\mu_0}{\mu_1}\frac{v_1}{u_1}\right)\left(\frac{\mu_1}{\mu_2}\frac{v_2}{u_2}\right)\cdots\left(\frac{\mu_{n-1}}{\mu_n}\frac{v_n}{u_n}\right)$$

or
$$m = \frac{\mu_0}{\mu_n}\frac{v_1 v_2 \ldots v_n}{u_1 u_2 \ldots u_n}.$$

If through P_0 we draw an arbitrary ray inclined at an angle α_0 to the axis, this inclination is successively changed by the refracting surfaces to $\alpha_1, \alpha_2, \ldots, \alpha_n$.

For each refraction Helmholtz's theorem holds:

$$\mu_0 y_0 \alpha_0 = \mu_1 y_1 \alpha_1,\ \mu_1 y_1 \alpha_1 = \mu_2 y_2 \alpha_2, \ldots \mu_{n-1} y_{n-1} \alpha_{n-1} = \mu_n y_n \alpha_n.$$

Hence it follows that Helmholtz's theorem holds for *any number of refractions*, so that we may write

$$\mu_0 y_0 \alpha_0 = \mu_n y_n \alpha_n$$

with complete generality.

By p. 79 the **angular magnification** or **convergence ratio** is the ratio of the inclinations of the rays at the end and at the beginning of the path of the ray, i.e. $k = \alpha_n/\alpha_0$, whereas the lateral magnification $m = y_n/y_0$. By the last form of Helmholtz's equation, therefore:

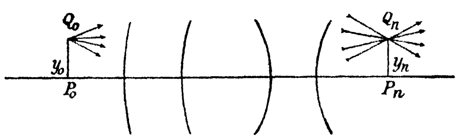


Fig. 23.—Principal points (P_0 and P_n) of an optical system

$$km = \frac{\mu_0}{\mu_n}$$

This equation expresses the deeper meaning of Helmholtz's theorem: *The product of the lateral magnification and the angular magnification depends only on the refractive indices of the first and last media and is independent of all the intermediate media.*

Principal Points and Principal Planes.* We see at once from fig. 22 and also from the expression for the lateral magnification m that in a system of coaxial spherical refracting surfaces there must be an object-point and a corresponding image-point such that the lateral magnification of an object at right angles to the axis is unity. If the object P_0Q_0 (y_0) is at one of these points (fig. 23) an upright image of the same size, P_nQ_n (y_n), appears at the corresponding point.

These special points are known as *principal points* of the system. All the rays which pass through one principal point must pass through the other principal point. This property, however, is not specially characteristic of the principal points, as it is possessed by any pair of corresponding points. If, however, we describe planes passing through the principal points at right angles to the axis (the *principal planes*), each point of one principal plane has for its image a point

* The introduction of the principal points and principal planes for any optical system of coaxial spherical surfaces is due to the mathematician K. F. Gauss (1840; see Vol. III, p. 195).

lying in an exactly similar position in the other principal plane. If the two principal planes were superposed on one another, the object-points and the image-points remaining fixed on them, the object-points and the image-points would exactly coincide.

If we know the principal planes of a system, we know that every ray passing through a point in one principal plane must pass through a point occupying exactly the same position in the second principal plane, and we do not need to trace the path of the ray through the system at all.

In fig. 24 the ray R is drawn to meet the first principal plane at A; it leaves the second principal plane at A′ as the ray R′. We may therefore indicate by the dotted line joining A and A′ and running parallel to the axis of the system that we do not wish to say anything about the course of the ray within the system. Of course the direction of R′ has not yet been determined; in the figure it has been drawn quite arbitrarily.

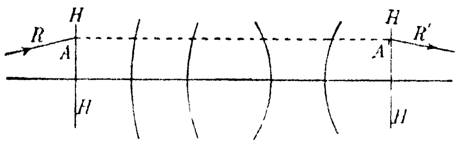


Fig. 24.—Principal planes (HH, H′H′) of an optical system

To determine the position of the principal points, we have to put $m = 1$ in the last expression for the lateral magnification and then use the general system of equations along with it to evaluate u_1 and u_n. Here we cannot give this calculation for the general case.

If we know both the principal planes HH and H′H′ and the foci F and F′ (defined as for a single lens) of a refracting system,* we can tell how rays which are parallel to the axis in the object space behave in the image space. If the ray S_1 parallel to the axis (fig. 25) meets the principal plane in the object space, or first principal plane, at A, it must leave H′H′, the principal plane in the image space, or second principal plane, at the corresponding point A′, and must then pass through F′, the second focus. The ray S_2 starting from the first focus F and meeting the first principal plane HH at B leaves the second principal plane at the corresponding point B′ as a ray S′ parallel to the axis. The two dotted lines AA′, BB′ parallel to the axis, however, do not represent the actual course of the rays through the system, but are merely geometrical construction lines joining corresponding points of the two principal planes.

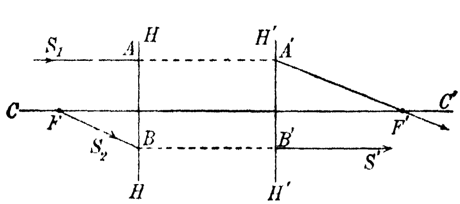


Fig. 25.—The principal planes, HH and H′H′, and the foci, F and F′

*[The foci and principal points are often referred to as the *cardinal points* of the refracting system (p. 85).]

Construction of the Image.—Let PR in fig. 26 be the object, perpendicular to the axis, of which an image is to be formed by the optical system. From the point Q of the object we then draw two rays, QA parallel to the axis and QF

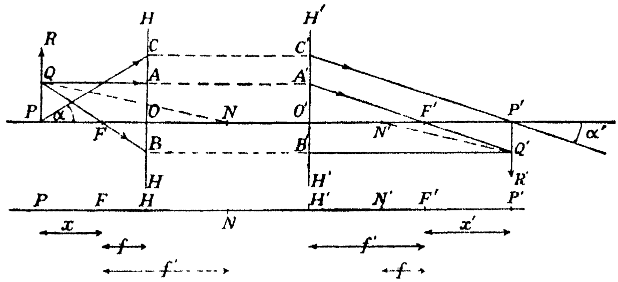


Fig. 26.—An object PQR and the image P′Q′R′ formed by an optical system; O and O′ are the principal points, HH and H′H′ the principal planes, F and F′ the foci, and N and N′ the nodal points.

through the first focus F. A and B, the points of intersection of these rays with the principal plane HH, correspond to the points A′, B′ of the principal plane H′H′, which are at the same distances from the axis as A and B. We then draw A′F′ through the second focus F′ and B′Q′ through B′ parallel to the axis. The two rays intersect at Q′, the image of Q. Other points of the object may be found in the same way, so that we finally obtain the image P′R′ perpendicular to the axis. This method of constructing the image was first used by J. B. Listing * (1851).

Nodal Points.—The convergence ratio $k = \alpha_n/\alpha_0$ may also take the special value unity for a definite position of the corresponding object- and image-points. This means that the inclinations of all the rays starting from the object-point are the same as the inclinations of the rays passing through the corresponding image-point.

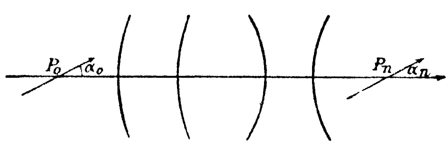


Fig. 27.—Nodal points (P_0 and P_n) of an optical system

The position of these special points, the so-called *nodal points*, may be calculated by putting $\alpha = \alpha'$ in Helmholtz's equation. We then obtain $fy = -f'y'$ for the nodal points. Now $m = y'/y = -x'/f' = -f/x$. Hence $x = f'$ and $x' = f$. That is, the nodal points N and N′ (fig. 26) lie at the distances f' and f to the right of the foci respectively. (In fig. 26 f is negative). If $\mu = \mu'$, i.e. if the optical media behind and in front of the optical system are the same, the focal lengths are equal and opposite. In this case the equation $fy = -f'y'$ for the nodal points becomes $y = y'$ and the *nodal points then coincide with the principal points.*

The characteristic property of the nodal points may also be expressed by the statement that all the rays passing through one nodal point leave the other nodal point parallel to their original direction (fig. 27).

* J. B. Listing (1808–1882) was born in Frankfurt and held a professorship at Göttingen.

6. Image formed by a Lens of Finite Thickness (Thick Lens).

Combination of Two Optical Systems to form a single Optical System.

(*a*) *Graphical Illustration and Relationships.*—Suppose that we are given the foci and principal planes of two optical systems with the same axis. Let the first system have the principal points H_1 and H_1' and the foci F_1 and F_1', and let the second system have the principal points H_2 and H_2' and the foci F_2 and F_2'. In fig. 28 the positions of the planes and points are shown diagrammatically; as the focal lengths are to be measured *positive* to the *right* from the principal points, it is arbitrarily assumed that both the foci are to the right of the principal planes of both systems. Let $H_1F_1 = f_1$ be the first focal length of the first system and $H_1'F_1' = f_1'$ the second; let $H_2F_2 = f_2$ be the first focal length of the second system and $H_2'F_2' = f_2'$ the second. Further, let the distance $H_1'H_2$ between

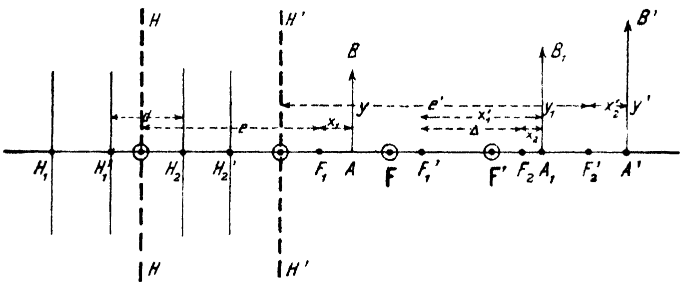


Fig. 28.—Combination of two optical systems with principal planes H_1, H_1' and H_2, H_2', and foci F_1, F_1' and F_2, F_2', to form a single optical system with principal planes HH, H'H' and foci F, F'. The first system forms the image A_1B_1 of the object AB, and the second forms the image A'B' of the first image.

the principal planes of the two systems which are "next" one another be d and the distance $F_1'F_2$ between the foci "next" one another (the so-called **optical interval**) be Δ. Now let there be an object AB (y) at the point A on the axis in the object space of the first system. The first system produces an image of AB, namely, A_1B_1 (y_1) at the point A_1 in the image space of the first system. This image is at the same time the object for the second system, which gives rise to the image A'B' (y') at A'. That is, the image of AB at A formed by the combined optical system is A'B' at A'. If in addition we denote the distances of the images and objects from the foci associated with them, measured positive to the right, as follows: $F_1A = x_1$, $F_1'A_1 = x_1'$, $F_2A_1 = x_2$, and $F_2'A' = x_2'$, we have

$$x_1x_1' = f_1f_1', \quad \ldots \ldots \ldots \quad (1)$$

$$x_2x_2' = f_2f_2'. \quad \ldots \ldots \ldots \quad (2)$$

Further, we see from the figure that

$$H_1'H_2 + H_2F_2 = H_1'F_1' + F_1'F_2$$

or

$$d + f_2 = f_1' + \Delta, \quad \ldots \ldots \ldots \quad (3)$$

and

$$F_1'A_1 = F_1'F_2 + F_2A_1$$

or

$$x_1' = \Delta + x_2. \quad \ldots \ldots \ldots \quad (4)$$

(*b*) *The Foci.*—The second focus F' of the combined system is the point where rays parallel to the axis in the object space of the system H_1H_1' are reunited.

For these rays $x_1 = \infty$, so that the corresponding value of x_1' is zero, by (1), $x_2 = -\Delta$, by (4), and $(x_2')_\infty = -f_2f_2'/\Delta$, by (2). Here $(x_2')_\infty$ means the distance of the second focus F′ of the combined system from the focus F_2'. Similarly, we obtain the position of the first focus of the combined system by putting $x_2' = \infty$ in (2) and finding the corresponding values of the other quantities. The equation (2) gives $x_2 = 0$; (4) gives $x_1' = \Delta$ and (1) gives $(x_1)_\infty = f_1f_1'/\Delta$. Here $(x_1)_\infty$ means the distance of the first focus F of the combined system from F_1.

(*c*) *The Principal Planes.*- To determine the positions of the principal planes of the combined system we must find the position of the two corresponding points A and A′ for which the magnification y'/y has the value unity. Now by p. 83

$$\frac{y_1}{y} = -\frac{f_1}{x_1} \text{ and } \frac{y'}{y_1} = -\frac{f_2}{x_2}$$

so that

$$\frac{y'}{y} = \frac{y'}{y_1}\frac{y_1}{y} = \frac{f_1f_2}{x_1x_2};$$

hence we must have

$$\frac{f_1f_2}{x_1x_2} = 1. \qquad (5)$$

Substituting for x_2 from (4), we have

$$f_1f_2 = x_1(x_1' - \Delta) = x_1x_1' - x_1\Delta$$

or, using (1),

$$x_1\Delta = f_1f_1' - f_1f_2,$$

and

$$e = x_1 = \frac{f_1(f_1' - f_2)}{\Delta},$$

where e denotes the distance of the first principal plane H of the combined system from the focus F_1.

The special value of x_2' corresponding to this is found by substituting in (1), (4), and (2). We have

$$x_1' = \frac{f_1f_1'}{x_1} = \frac{f_1'\Delta}{f_1' - f_2},$$

$$x_2 = x_1' - \Delta = \frac{f_1'\Delta - f_1'\Delta + f_2\Delta}{f_1' - f_2} = \frac{f_2\Delta}{f_1' - f_2},$$

and

$$e' = x_2' = \frac{f_2f_2'}{x_2} = \frac{f_2'(f_1' - f_2)}{\Delta},$$

where e' denotes the distance of the second principal plane H′ from F_2'.

(*d*) *The Focal Lengths.*—The focal lengths of the combined system are the distances of the foci from the corresponding principal planes H and H′. We have (see above)

$$f = \text{HF} = \text{F}_1\text{F} - \text{F}_1\text{H} = (x_1)_\infty - e = \frac{f_1f_1'}{\Delta} - \frac{f_1(f_1' - f_2)}{\Delta} = \frac{f_1f_2}{\Delta},$$

$$f' = \text{H}'\text{F}' = \text{F}_2'\text{F}' - \text{F}_2'\text{H}' = (x_2')_\infty - e' = -\frac{f_2f_2'}{\Delta} - \frac{f_2'(f_1' - f_2)}{\Delta} = -\frac{f_2'f_1'}{\Delta}.$$

We have now determined the **cardinal points** H, H′, F, F′ of the combined system and also its focal lengths. If we denote the distances of the object A and the image A′ from the foci F and F′ of the combined system respectively by $\text{FA} = x$, $\text{F}'\text{A}' = x'$, we again have the equation

$$xx' = ff'$$

connecting the positions of the object and the image.

Focal Lengths and Cardinal Points of a Lens of Finite Thickness.

(*a*) *Introductory Remarks.*—In fig. 29 L_1L_1, L_2L_2 denote two spherical surfaces with centres of curvature M_1, M_2, radii of curvature r_1, r_2, and poles S_1, S_2. Let S_1S_2, the distance between the two poles, be d. Let the medium (e.g. glass) betweeen the two spherical surfaces have the refractive index μ, and let the refractive index of the medium (e.g. air) outside these spherical surfaces be unity.

Each of the two refracting surfaces L_1L_1 and L_2L_2 may then be regarded as one of the individual optical systems of the previous discussion. The principal planes of a single spherical refracting surface coincide with the surface itself.* H_1 and H_1' of the previous argument coincide with S_1, H_2 and H_2' with S_2.

(*b*) *The Focal Lengths.*—Let P_0 ($=$P) be the object-point on the axis for the first refracting surface, P_1 its image-point, which acts as object for the surface L_2L_2, and finally let P′ ($=P_2$) be the image formed by the lens of the object P_0 ($=$P). If we reckon the distances of these points from the poles S_1, S_2 (positive to the

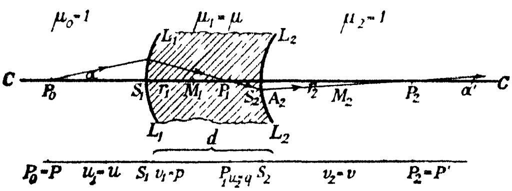


Fig. 29.—A thick lens

right) and denote S_1P_0 ($=S_1$P) by u, S_1P_1 by p, S_2P_1 by q, and S_2P_2 ($=S_2$P′) by v, and if we further put $\mu_0 = 1$, $\mu_1 = \mu$, $\mu_2 = 1$, the fundamental equation for a spherical surface (p. 65) gives

$$\frac{\mu}{p} = \frac{1}{u} + \frac{\mu - 1}{r_1},$$

$$\frac{1}{v} = \frac{\mu}{q} + \frac{1 - \mu}{r_2}.$$

From these we obtain the focal lengths corresponding to the surfaces L_1L_1, L_2L_2 by making u, p, q, v successively infinite and evaluating the corresponding points. Using the notation of the previous section (see above) we then obtain

$$f_1 = u_\infty = -\frac{r_1}{\mu - 1};\ f_1' = p_\infty = \frac{\mu r_1}{\mu - 1};$$

$$f_2 = q_\infty = -\frac{\mu r_2}{1 - \mu} = +\frac{\mu r_2}{\mu - 1};\ f_2' = v_\infty = \frac{r_2}{1 - \mu} = -\frac{r_2}{\mu - 1}.$$

Knowing the value of d and the values of the individual focal lengths f_1, f_1' f_2, f_2' obtained from r_1, r_2, and μ, we can immediately calculate Δ, the distance between the foci next each other (the *optical interval*, in the microscope the *optical tube-length*), from equation (3) of the previous section, $\Delta = d + f_2 - f_1'$.

* For by the fundamental equation (p. 65) we have $\frac{\mu'}{v} = \frac{\mu}{u} + \frac{\mu' - \mu}{r}$, and by fig. 20 (p. 77) $m = \frac{y'}{y} = \frac{v - r}{u - r}$. For given values of m, μ', μ, and r, these two equations determine u and v. If in particular $m = 1$, u and v must be equal. This is true if either $u = v = r$, or $u = v = 0$. That is, the principal planes coincide in the tangent plane, to the surface at S.

(c) *The Principal Points.*—To fix the positions of the principal points of our lens relative to the poles of the lens, we have to add to the values e and e' of the previous section (see above), giving the distances of the principal planes from the foci F_1 and F_2', the distances S_1F_1, S_2F_2' of these latter points from the poles S_1 and S_2. As, however, the principal points of the two refracting surfaces coincide with these poles, the first and second focal lengths are $S_1F_1 = f_1$ and $S_2F_2' = f_2'$. Let the distances of the principal points of the lens as a whole from the poles be $S_1H = s_1$, $S_2H' = s_2$. Then

$$s_1 = S_1H = S_1F_1 + F_1H = f_1 + e = f_1 + \frac{f_1(f_1' - f_2)}{\Delta} = \frac{f_1(\Delta + f_1' - f_2)}{\Delta} = \frac{f_1 d}{\Delta},$$

$$s_2 = S_2H' = S_2F_2' + F_2'H' = f_2' + e' = f_2' + \frac{f_2'(f_1' - f_2)}{\Delta}$$

$$= \frac{f_2'(\Delta + f_1' - f_2)}{\Delta} = \frac{f_2'd}{\Delta}.$$

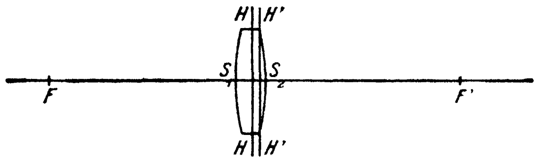


Fig. 30.—A glass lens ($\mu = 1 \cdot 5$, $r_1 = 4$ cm., $r_2 = -2$ cm.) with its principal planes and foci

(d) *The Focal Lengths of the Lens as a whole.*—From the results given in last section (see above) we find that the focal lengths of the lens are

$$f = \frac{f_1 f_2}{\Delta}, \quad f' = -\frac{f_2' f_1'}{\Delta}.$$

As the refracting medium is the same behind and in front of the lens (namely, air), f must be equal to $-f'$, by the general theorem $\mu f + \mu' f' = 0$ (p. 76). If we calculated the values of f_1, f_1', f_2, f_2' from the quantities r_1, r_2, d, and μ, and substituted these values for Δ and also in those for s_1, s_2, f, and f', we should obtain final formulæ for these quantities involving the given quantities only. We shall not write down these formulæ, however, as they are fairly complicated and less easy to grasp than the equations given above.

We shall, however, consider the case of an infinitely thin lens. Here $d = 0$, i.e. $\Delta = f_2 - f_1'$ (see above) and

$$f' = -\frac{f_2' f_1'}{f_2 - f_1'} \text{ or } \frac{1}{f'} = -\frac{1}{f} = -\frac{f_2 - f_1'}{f_2' f_1'} = -\frac{f_2}{f_2' f_1'} + \frac{1}{f_2'}.$$

By substitution (see above) we obtain

$$\frac{1}{f'} = -\frac{1}{f} = +\frac{\mu r_2(\mu - 1)^2}{(\mu - 1) r_2 \mu r_1} - \frac{\mu - 1}{r_2} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

for the refractive power of the lens, in agreement with p. 67.

(e) *The Distance between the Principal Planes.*—If the thickness of the lens is d, this distance is

$$u = d - s_1 + s_2 = \frac{d}{\Delta}(\Delta - f_1 + f_2')$$

by the foregoing.

(*f*) *Numerical Examples.*—We shall illustrate the general results we have just obtained by means of a few numerical examples. We first consider a biconvex lens of glass with the following dimensions (fig. 30):

$$r_1 = 4 \text{ cm.}, \ r_2 = -2 \text{ cm.}, \ d = 0{\cdot}4 \text{ cm.}, \ \mu = 1{\cdot}5.$$

Then
$$f_1 = -\frac{r_1}{\mu - 1} = -8 \text{ cm.}, f_2 = \frac{\mu r_2}{\mu - 1} = -6 \text{ cm.},$$

$$f_1' = \frac{\mu r_1}{\mu - 1} = +12 \text{ cm.}, f_2' = -\frac{r_2}{\mu - 1} = +4 \text{ cm.},$$

$$\Delta = d + f_2 - f_1' = 0{\cdot}4 - 6 - 12 = -17{\cdot}6 \text{ cm.},$$

$$s_1 = \frac{f_1 d}{\Delta} = \frac{-8 \times 0{\cdot}4}{-17{\cdot}6} = \frac{2}{11} = 0{\cdot}18 \text{ cm.}, \ s_2 = \frac{f_2' d}{\Delta} = \frac{4 \times 0{\cdot}4}{-17{\cdot}6} = -\frac{1}{11} = -0{\cdot}09 \text{cm.},$$

$$f = \frac{f_1 f_2}{\Delta} = \frac{-8(-6)}{-17{\cdot}6} = -\frac{48}{17{\cdot}6} = \frac{-30}{11} = -2{\cdot}73 \text{ cm.},$$

$$f' = -\frac{f_2' f_1'}{\Delta} = \frac{-4 \times 12}{-17{\cdot}6} = +\frac{30}{11} = +2{\cdot}73 \text{ cm.}$$

The distance between the principal planes is

$$u = d - s_1 + s_2 = 0{\cdot}4 - 0{\cdot}18 - 0{\cdot}09 = +0{\cdot}13 \text{ cm.}$$

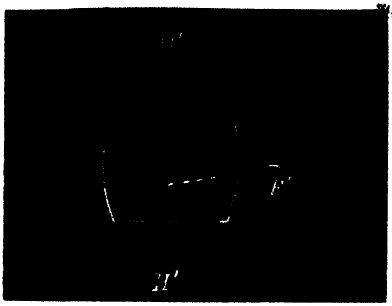
As the equations for Δ, s_1, s_2, f, and f' show, these quantities vary considerably with d, the thickness of the lens. If $d = 6$ cm., so that the centres of curvature coincide, $u = 0$, so that the two principal planes also coincide. If $d = 8$ cm., we have

$$\Delta = -10 \text{ cm.}, \ s_1 = +6{\cdot}4 \text{ cm.}, \ s_2 = -3{\cdot}2 \text{ cm.},$$
$$f' = -f = 4{\cdot}8 \text{ cm.}, \ u = -1{\cdot}6 \text{ cm.}$$

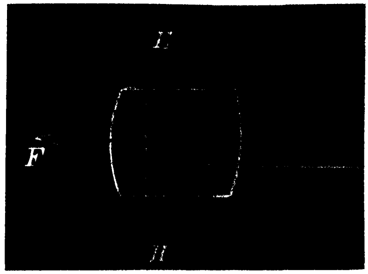
The principal planes have now exchanged places and have moved to a distance 1·6 cm. apart, and the focal length has considerably increased. If $d = 18$ cm., $\Delta = 0$ and the principal planes and foci have moved to an infinite distance; the lens no longer has the properties of a convergent lens. (It now forms an image like a telescope; p. 125). If $d > 18$ cm., Δ becomes positive and s_1, s_2, f, and f' change their signs. Then the principal planes lie outside the lens in their original order and the lens becomes a divergent lens (although it is thicker in the centre than at the edges).

Optical Systems formed by the Combination of Several Thin Lenses.—Let an optical system consist of two lenses which are so thin that we may take the principal planes of the lenses as coinciding with the lenses (when $d = 0$ we have $s_1 = s_2 = u = 0$ by (*c*) and (*e*)). Further let the medium behind and in front of the two lenses be the same, namely, air. Then $-f_1 = f_1'$ and $f_2' = -f_2$. The distance d' between the two principal planes next one another is now merely the distance between the two thin lenses; the distances s_1 and s_2 are always to be measured from the lenses. Under these conditions our formulæ simplify to

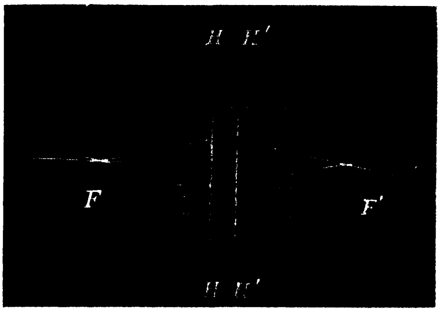
$$\Delta = d' - f_1' - f_2', \ s_1 = -\frac{f_1' d'}{\Delta}, \ s_2 = \frac{f_2' d'}{\Delta},$$
$$-f' = +f = \frac{f_1' f_2'}{\Delta} = +\frac{f_1 f_2}{\Delta}.$$



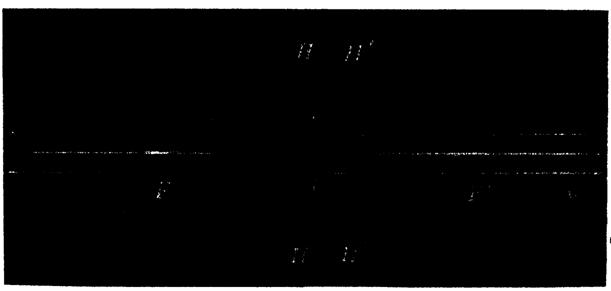
Ch. IV. Fig. 34. Second principal plane and focus of a thick lens



Ch. IV. Fig. 35. First principal plane and focus of a thick lens



Ch. IV. Fig. 36. The two principal planes and foci



Ch. IV, Fig. 37. Principal planes and foci of a concavo-convex lens

The question whether the system acts as a convergent lens or a divergent lens depends on the sign of Δ (see paragraph (f) above).

The last equation is often written in another form, involving the refractive power of the system instead of the focal length:

$$\frac{1}{f'}=-\frac{\Delta}{f_1'f_2'}=-\frac{d'-f_1'-f_2'}{f_1'f_2'}=\frac{1}{f_1'}+\frac{1}{f_2'}-\frac{d'}{f_1'f_2'}.$$

If the two thin lenses are actually touching one another, so that d' may be put equal to zero, we have

$$\frac{1}{f}=\frac{1}{f_1}+\frac{1}{f_2}.$$

The powers of two thin lenses in contact (measured in dioptres) are additive.

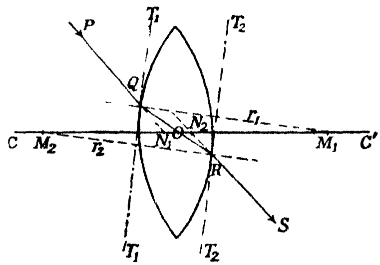


Fig. 31.—Nodal points (N_1 and N_2) and optical centre of a thick lens

Nodal Points.—The position of the nodal points of a thick lens may also be determined by the following considerations (fig. 31). M_1, M_2 are the centres of curvature and r_1, r_2 the radii of curvature of the two spherical surfaces by which the lens is bounded. Through M_1 and M_2 we draw two parallel radii M_1Q and M_2R, and join QR. Let this represent that part of a certain ray of light which is included within the lens. If we draw the tangent planes at Q and R to the spherical surfaces, T_1T_1 and T_2T_2, these form the boundaries of a flat plate, and from the refractive index we can calculate the path of the ray of light P which enters the plate at Q and leaves it at R as the parallel but laterally displaced ray S. By producing PQ and SR we obtain their intersections (N_1 and N_2) with the axis of the lens CC′. These two points are the nodal points of the lens. As the medium is the same on both sides of the lens (namely, air) the nodal points coincide with the principal points (p. 83). The point O where the ray QR cuts the axis of the lens is called the **optical centre** of the lens.

From the similar triangles OQM_1, ORM_2 we have

$$\frac{OQ}{OR}=\frac{r_1}{r_2};$$

hence

The optical centre of a lens divides the thickness of the lens in the ratio of the radii of curvature.

Lecture Experiments.

(*a*) *Parallel Displacement.*—The parallel displacement of a ray of light by a very thick lens may be demonstrated objectively by letting a ray of light pass obliquely through the lens. Fig. 32 (Plate IV) is reproduced from a photograph, the ray being made visible by blowing tobacco smoke into its path.

The thick lenses used in the experiments illustrated in figs. 33–37 (Plates IV, V) were made by cementing plano-convex lenses to opposite faces of a glass cube of side 4 cm. by means of Canada balsam. The paths of the rays were made visible by tobacco smoke and were then photographed, while the outlines of the lens and the dotted construction lines were added to the photographs subsequently.

Fig. 33 shows that a ray which falls on the lens in such a way that it leaves

the lens parallel to its former direction but laterally displaced meets the axis of the lens in two points; these are the nodal points N and N′ (cf. p. 83).

(*b*) *The Principal Planes.*—In fig. 34 (Plate V) three parallel light rays are shown falling on a biconvex lens with different curvatures on the two sides; these rays are made to converge by the lens and intersect at the second focus F′. If we produce the incident parallel rays forwards and the rays leaving the lens backwards, they will intersect. The plane H′H′ passing through these intersections at right angles to the axis of the lens is the second principal plane; its distance from the second focus is the second focal length.

In fig. 35 (Plate V) three rays starting from the point F are made parallel on passing through the lens. The starting-point of these rays is the first focus F. If we produce the incident rays forwards and the emergent rays backwards, these prolongations intersect in the first principal plane HH.

In the experiment shown in fig. 36 (Plate V) three parallel rays fall on the same lens from either side. We thus obtain the two foci F and F′ simultaneously. The distances of the foci from the corresponding principal planes are the two focal lengths; here, as the rays are passing from air to glass and back to air, these are equal, although the distances of the foci from the poles of the lens are different.

The unsymmetrical position of the foci and principal planes is shown in a particularly striking way by the experiment of fig. 37 (Plate V). Here the glass cube has a plano-convex lens cemented on to one side and a concavo-plane lens on to the other. The experiment makes it clear that the principal plane H′H′ lies entirely outside the lens. Accordingly the focus F is close to one pole, while the focus F′ is a great deal farther from the other pole; yet the two focal lengths are the same.

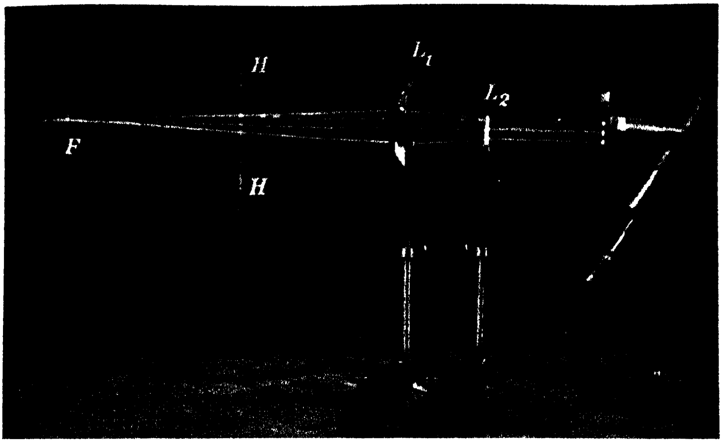
In figs. 38, 39 (Plate VI) three parallel rays fall on a combination of two lenses, of which L_1 is a convexo-plane lens and L_2 a concavo-plane lens. In this system both the principal planes HH, H′H′ lie outside the system on the same side of it. The unsymmetrical position of the foci is clearly shown; yet here again the two focal lengths are the same, as the first medium and the last medium have the same refractive index.

7. Image, formed by a Thick Lens, of an Object Perpendicular to its Axis.

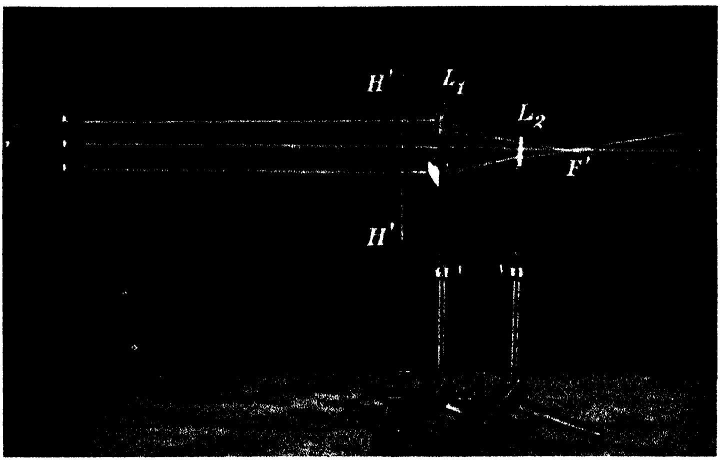
In the experiment of fig. 40 (Plate VII) three rays start from a point Q not lying on the axis of the lens. One ray goes through the first focus F; it leaves the lens parallel to the axis. If we produce the incident and emergent rays, they meet in the first principal plane HH. The second ray enters the lens parallel to the axis; on leaving the lens it is inclined to the axis and passes through the second focus F′. The prolongations of the incident and emergent rays meet in the second principal plane H′H′. The third ray is in the direction of the first nodal point, which in this case coincides with the first principal point; it leaves the lens parallel to its original direction, but displaced in such a way that it appears to come from the second nodal point, which coincides with the second principal point. The three selected rays meet again at the point Q′ corresponding to Q.

This experiment is arranged on the plan of fig. 26, p. 83, in which the image of an object was obtained by means of special rays when the principal points and foci of the lens were known.

If the first medium, i.e. the medium in which the object is situated, has a refractive index which differs from that of the last medium, as, e.g., in the eye, the principal points do *not* coincide with the nodal points. In this case, therefore, the positions of the foci, the principal points, and the nodal points must be known before we can construct the image of a given point by means of particular rays.



Ch IV, Fig 38 First principal plane and focus of a combination of two lenses



Ch IV, Fig 39 Second principal plane and focus of a combination of two lenses

CHAPTER V

Geometrical Optics: Aberration

1. Aberration.

In discussing the properties of lenses, we have hitherto assumed that points not on the axis lie so near to it that the rays of light starting from them meet the axis at an angle so small that sine, tangent, and angle are interchangeable. If we dispense with this limitation, the mathematical relationships become very complicated. Hence we shall confine ourselves to experimental investigation of the new phenomena which occur. The figures have been obtained by photographing the paths of the rays as before.

In the experiment of fig. 1 (Plate VII) a ray of light passes as an axial ray through the centre of a plano-convex lens with its plane surface facing the incident light. At a short distance from it two rays travel parallel to the axis; these are made to converge by the lens and meet the axis behind the lens at F_m. This point is the second focus, in the sense in which the term was used in the previous discussion. There are also two rays parallel to the axis at a greater distance from it which are incident on regions of the lens near its edge (so-called **marginal rays**). These are likewise made to converge by the lens and then cut the axis at a point F_r. This point, however, does not coincide with F_m, but lies considerably nearer the lens. If, then, we imagine a whole pencil of parallel rays incident on the lens from the left, all the rays of the pencil will no longer be reunited at a single point, the rays from each **zone** * having the same meeting-point, which, however, differs from that of every other zone.

The distance F_mF_r is called the **longitudinal aberration.**

The phenomenon shown in fig. 1 is partly due to the spherical form of the refracting surfaces, and is commonly referred to as **spherical aberration.** It is possible, it is true, to calculate the forms of lenses, and to produce actual lenses of the forms calculated, which will reunite all the rays from a given object-point at a *single* point, i.e. without aberration. But even in this case aberration will occur for any other object-point. Aberration is by no means a phenomenon peculiar to

* Here the word zone is used in the mathematical sense of spherical zone; it means a narrow strip of the lens surface which is all at the same distance from the axis.

the spherical surface. In actual fact, the aberration is smaller with lenses bounded by spherical surfaces than for other forms of lenses, unless we are only seeking to recombine those rays which have a specially selected divergence.

Further, aberration is not only a function of the lens form, but also a function of the angle at which the light rays meet the surfaces of the lens. This we see by the following experiment.

We turn round the plano-convex lens used in the previous experiment, so that its curved surface is directed towards the incident pencil of parallel rays. We then have the phenomenon shown in fig. 2 (Plate VII); F_m and F_r move very close together, so that the aberration is now considerably less than it was in the experiment of fig. 1.

Aberration also occurs with concave lenses. To demonstrate the existence of aberration in this case, of course, we have to find the points of intersection with the axis by backward prolongation of the rays.

The aberration can be got rid of for certain definite zones by suitable combination of two or more lenses. A system of this kind is said to be *corrected* for spherical aberration. Such a system may be free of aberration for two or at most three zones, never for all zones. For points on the axis, however, it is possible to produce systems which are satisfactorily devoid of aberration.

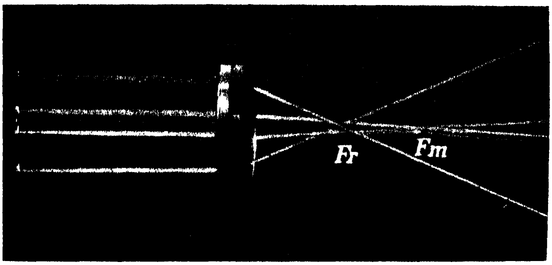
Just as the focal distances vary for the different zones of a lens, so do the focal distances for one and the same zone vary for light of different *colours*. For this so-called **chromatic aberration** see p. 167.

2. Astigmatism.

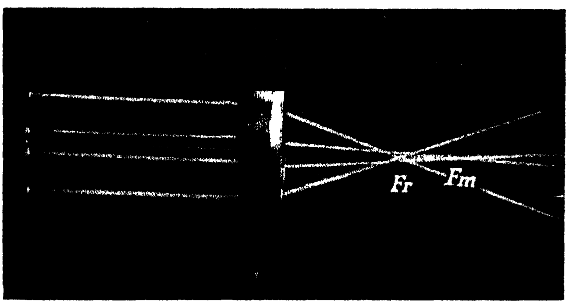
If a lens bounded on one side by a portion of a spherical surface and on the other by a portion of a cylindrical surface is set up in such a way that the straight lines generating the cylindrical surface stand vertical, a vertical section passing through the axis is identical with the principal section of a plano-convex lens and a horizontal section is identical with the principal section of a biconvex lens, the curvature of the one surface of the biconvex lens being equal to the curvature of the convex surface of the plano-convex lens. It follows that (if we neglect the longitudinal aberration) the lens will cause parallel rays lying in a *vertical* plane and parallel rays lying in a *horizontal* plane to meet at different points, the latter point lying nearer the lens than the former.

This experiment is shown in fig. 3 (Plate VIII). Parallel rays pass through a stop in the form of a circular ring and fall on a lens which is cylindrical on the left (the lines generating the cylinder being vertical) and spherical on the right. In the lower half of the figure the observer is looking along a horizontal plane, in which the rays appear to meet at the point F_h. In order that the observer may simultaneously gain a view of the same rays from above, i.e. in the vertical

Ch IX, Fig 40 Image of an object at right angles to the axis of a thick lens



Ch V, Fig 1 The focal length for marginal rays is less than that for axial rays



Ch V, Fig 2 Diminution of the aberration when the lens is turned round

plane, a long narrow mirror is set up above the path of the rays at an angle of 45°. Here we see the point F_v where the rays unite in a vertical plane. There is no point in which all the rays in any inclined plane are united. If we describe planes at right angles to the lens at varying distances from the lens, we in general obtain elliptic sections; at the points F_h and F_v the vertical and horizontal axes of the ellipses respectively vanish, giving rise to a horizontal or vertical straight line.

The two middle parts of fig. 3 were obtained directly by photography of the pencil of rays made visible by tobacco smoke. (The mirror does not come out in the photograph.) The two figures above and below these were obtained by putting small, slightly dusty pieces of plate glass in the path of the same pencil of rays, so as to make the cross-sections of the pencil clearly visible. These sections were then photographed somewhat obliquely from the front. Hence the two parts of the figure which represent the cross-sections of the pencil must be rotated through about 45° round a vertical axis if they are to correspond exactly to the two middle parts of the figure; then, however, all the sections would of course appear as short straight lines and their shapes could not be seen.

The above is also true for rays diverging from a point; hence the lens of fig. 3 is incapable of forming a *point image* of a point. For this reason the phenomenon which we have just described is called **astigmatism,*** and a lens which exhibits astigmatism is said to be astigmatic. The phenomenon is not uncommon in the case of the eye.

Astigmatism also occurs in a lens bounded on both sides by spherical surfaces if the incident rays are sharply inclined to the axis. To distinguish this case from the preceding we refer to it as the **astigmatism of oblique pencils.**

In fig. 4 (Plate VIII) a parallel pencil passing through a stop consisting of a ring of holes is made to traverse a biconvex lens with spherical surfaces, but the lens is rotated about a vertical straight line so that the axis of the lens makes an angle of 45° with the direction of the rays of light. As in the former experiment, the figure is derived from a photograph of the path of the rays. A mirror set up at an angle above the path of the rays enables the phenomenon to be observed in a vertical plane at the same time. The phenomenon is essentially the same as that shown in fig. 3: the effect of the lens on a ray inclined at a large angle to the axis is as if the lens were more strongly curved.

It follows that a lens cannot reunite in one point those rays which start from an object-point remote from the axis. The image is no longer *unique.*

3. The Sine Condition.

Owing to aberration and the astigmatism of oblique pencils, a point object in general does not give rise to a perfectly clear point image. On p. 92, however, we saw that by combination of several lenses it is possible to produce a corrected system which in practice

* Gr., *stigma,* a point; *astigmatic* accordingly means *without a point.*

gives rise to no aberration for points on the *axis*. That is, corrected systems can be constructed in which an axial point object gives rise to an axial image without aberration.

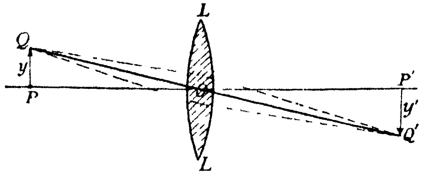


Fig. 5.—A thin pencil of light rays starting from the object Q is brought to a point at Q′ by the part of the optical system near the axis.

Let LL in figs. 5 and 6 denote an optical system without aberration for points on the *axis*; here, for the sake of simplicity, we shall represent the system by a single convex lens. Let P′ be the point on the axis which corresponds to the point P on the axis. If we imagine a *small* object PQ (y) placed at P at right angles to the axis, the system gives rise to the image P′Q′ (y') at right angles to the axis. The magnification $m = y'/y$ may be calculated by means of the equations developed in § 4 of last chapter (p. 77). These, however, are not strictly true except for a thread-like region along the axis. In such a case the rays do come to a single point; all the rays starting from Q meet again at Q′.

For an optical system with considerable *lateral extension*, however, the rays in general are not reunited in a single point. A narrow pencil of rays near the axis, with the central ray QO, meets at a point (Q′) different from that (Q_1') where a narrow pencil, with the central ray QR, which is refracted by the marginal portions of the system (figs. 5, 6) is reunited; for the pencil passing through the optical centre of the system (fig. 5) gives an image P′Q′ of magnitude y', whereas the pencil passing through the marginal region R (fig. 6) gives an image $P'Q_1'$ of magnitude y_1'.

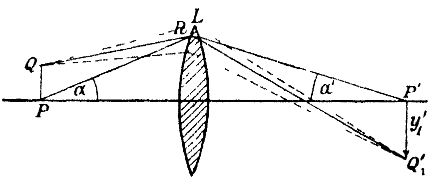
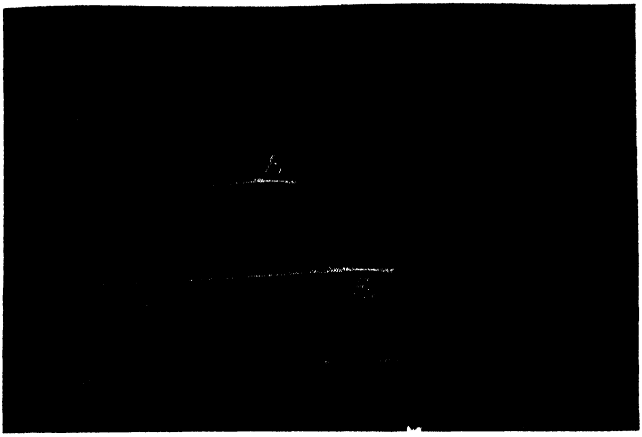


Fig. 6.—A thin pencil of light rays starting from the point Q is brought to another point Q_1' by the marginal zone of the optical system.

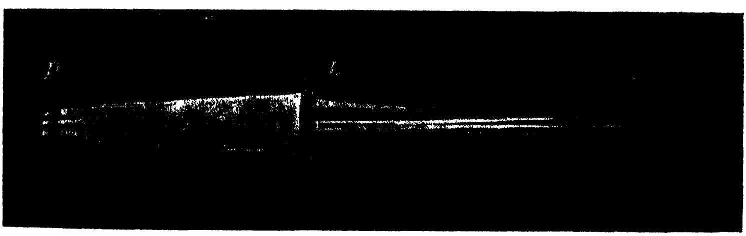
Suppose that a ray passing from P to R leaves the axis at the angle α, and that the refracted ray RP′ reaches it at the angle α'. If the angles α and α' are small, Helmholtz's equation is true (§ 4 of last chapter, p. 77). As the refractive index of the last medium is the same as that of the first, this equation becomes $m = y'/y = \alpha/\alpha'$, so that the magnification is independent of the inclination of the ray to the axis. This, however, is no longer true once α exceeds a definite small value. The magnification of the image therefore varies according to the region of the lens traversed by the rays producing it. In an actual case an axial point P may be very accurately reproduced as a point (e.g. the aplanatic points of a spherical surface (p. 105)), whereas the images Q′ and Q_1' of a point *near* the axis produced by different zones no longer coincide. Then the image of the point Q near the axis produced by the whole pencil of rays of aperture 2α becomes blurred.

Sharp images would, of course, be obtained if Helmholtz's equation could be validly extended, as a result of some peculiarity of the optical system, to rays inclined at a considerable angle to the axis. This, however, is not the case; the condition is inseparably associated with small values of the inclinations of the rays to the axis. In tech-



Ch V Fig 3 Astigmatism the focal distances in the two principal sections

Ch V, Fig 4 Astigmatism of oblique pencils



Ch V, Fig 8 Blurred image of points near the axis

nical optics it is essential to get rid of this limitation; for the greater the aperture of the pencil of rays, the greater the quantity of light and hence the brighter the images produced. Hence designers of optical instruments are keenly interested to know under what conditions a magnification of the image which is independent of the zone of the lens traversed can be obtained with pencils of finite aperture. This problem was first elucidated by HELMHOLTZ and more especially by ABBE. The condition which must be satisfied is the so-called **sine condition:**

$$\frac{\mu \sin\alpha}{\mu' \sin\alpha'} = \frac{y'}{y} = \text{const.}$$

Here μ, μ' denote the refractive indices in the object space and the image space and $y'/y = m$ is the lateral magnification. In addition to being the condition that the magnification shall be independent of the zone giving rise to the image, the sine condition is also the condition that an element of surface near the axis is reproduced without distortion by pencils of large aperture. ABBE deduced the sine condition from the wave theory of light. As HELMHOLTZ has shown, the requirement that a small element of surface near the axis shall give rise to the brightest possible image leads to the sine condition also.

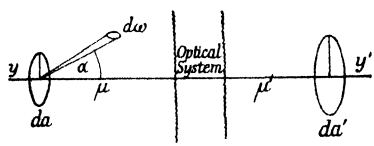


Fig. 7.—To illustrate the derivation of the sine condition

The sine condition may readily be deduced from the principle of energy. In fig. 7, da is an element of area at right angles to the axis of the (arbitrary) system, which is reproduced as the element of area da' which is also at right angles to the axis. Let the angle between a ray starting from da and the axis be α and the angle between the ray incident on da' and the axis be α'. If S is the energy density radiated vertically by unit area of da, and S′ that incident vertically on unit area of da', and $d\Phi$ the energy radiated per solid angle $d\omega$ in the direction making an angle α with the axis, we have

$$d\Phi = S\,da\,.\cos\alpha\,d\omega.$$

If da' is a sharp image which is to be as bright as possible, the whole of the radiation from da must be brought together again at da'. Hence we must have

$$d\Phi = S'da'\,.\cos\alpha'd\omega'.$$

The solid angle $d\omega$ may, as we know from spherical trigonometry, be replaced by $\sin\alpha\,d\alpha\,d\varphi$, where φ denotes the (geographical) longitude on the unit sphere. The whole luminous flux radiated up to the maximum angle α is therefore the sum of all the elementary luminous fluxes $d\Phi$ radiated in this solid angle, the summation being taken from 0 to α for the inclination to the axis and from 0 to 2π for φ. As in addition, we have $\cos\alpha\,d\alpha = d(\sin\alpha)$,

$$\Phi = S\,da\int_0^\alpha\int_0^{2\pi}\cos\alpha\,\sin\alpha\,d\alpha\,d\varphi$$

$$= \mathrm{S}\,da \int_0^{\sin\alpha} \int_0^{2\pi} \sin\alpha\, d(\sin\alpha)\, d\varphi$$
$$= \mathrm{S}\,da \,.\, \tfrac{1}{2}\sin^2\alpha \,.\, 2\pi.$$

By the above Φ must have the same value for da'. Hence

$$\mathrm{S}\, d\alpha \sin^2\alpha = \mathrm{S}' d\alpha' \sin^2\alpha'.$$

Now $\mathrm{S}/\mathrm{S}' = \mu^2/\mu'^2$. (This may be seen as follows: the energy-density in the first medium is $\mathrm{S} = \mathrm{DEH}$, in the second $\mathrm{S}' = \mathrm{D'EH}$, as the permeability may be taken as unity for both media; hence $\mathrm{S}/\mathrm{S}' = \mathrm{D}/\mathrm{D}' = \mu^2/\mu'^2$.) It follows that

$$\mu^2 \sin^2\alpha\, da = \mu'^2 \sin^2\alpha'\, da';$$

hence

$$\mu \sin\alpha \sqrt{da} = \mu' \sin\alpha' \sqrt{da'}.$$

As $\sqrt{(da'/da)}$ is the lateral magnification, we finally obtain

$$\frac{\mu \sin\alpha}{\mu' \sin\alpha'} = m.$$

This, then, is the condition that the energy emitted by a small element of area da is again concentrated on a small element of area da' of magnitude such that $da'/da = m^2$.

Points for which the sine condition holds and which at the same time are the images of one another devoid of longitudinal aberration were called aplanatic points * by ABBE.

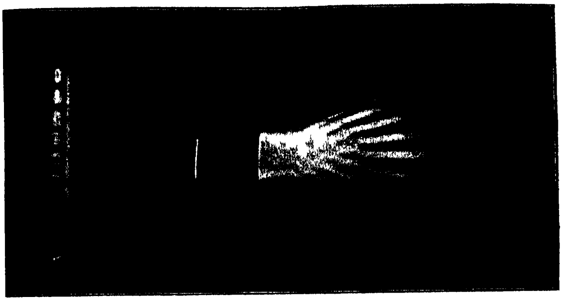
Helmholtz's equation must hold for geometrical reasons whenever straight rays give rise to a *point* image of constant magnification (p. 77), i.e. we must have $\mu' \tan\alpha'/\mu \tan\alpha = y/y' = \text{const.}$; the sine condition is inconsistent with this unless the inclinations are small. It follows that the production of point images of a region of arbitrary size by means of a pencil of wide aperture is *physically* impossible. If the sine condition is satisfied for a single pair of points on the axis, one point being the unique and distinct point image of the other (see the aplanatic points of a spherical surface (p. 105)), there can never be another pair of points on the axis for which the sine condition is satisfied at the same time.

Fig. 8 (Plate VIII) is from a photograph in which an image of three luminous object points P is formed by an aberrationless lens L. A pencil of rays passing through the centre of the lens and two pencils of rays passing through the lens near its edge are selected by means of a stop fixed to the lens. If we construct an image plane perpendicular to the axis at the point P′ where the image of the point on the axis is situated, we see that the images of points near the axis formed by the central rays are farther from the axis than the images formed by the marginal rays. In the figure this difference is, of course, very small, but it would give rise to a blurring of the images of points near the axis.

4. The Variation of Distortion with the Position of the Stop.

An ordinary simple lens exhibits both aberration and astigmatism; hence it gives rise to a blurred image, especially when used to produce an image of an extended object whose plane is at right angles to the axis of the lens, on a plane screen also at right angles to the axis.

* [Sometimes the term *aplanatic foci* is used.]

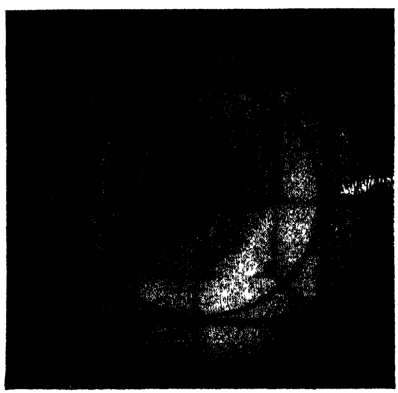


Ch. X. Fig. 9. A pencil of rays passing through a lens without a stop; the images lie on a curved surface.

Ch. X. Fig. 10. The centre and the marginal regions give sharp images at different distances from the lens; further, the radii and the arcs orthogonal to them give sharp images at different distances.

Ch. X. Fig. 11. The effect of aberrations; the image screen is set up in a position nearer the object than that corresponding to a sharp centre.

(Berek, *Grundlagen der praktischen Optik* (W. de Gruyter and Co., Berlin).)



In the experiment of fig. 9 (Plate IX) the object consists of a set of brightly illuminated points set up on one side of an ordinary biconvex lens * at right angles to the axis. Each of these points sends rays of light through the lens, which causes every separate pencil of rays to converge.

Fig. 9 at once shows that the regions where the individual pencils are most nearly reunited do not lie in a plane at right angles to the axis of the lens, but on a curved surface, the concave side of which faces towards the lens.

If we now place a plane vertical screen in the image space at right angles to the axis of the lens, only a small region of the image is sharp, and in fact the screen must be moved nearer the lens the greater the distance between the axis and the portion of the image which it is desired to render sharp. Here by sharpness we do not, of course, mean absolute sharpness, i.e. the representation of a point by a point, as the sharpness of all parts is impaired by the aberration and the non-fulfilment of the sine condition, and the sharpness of the marginal parts by the astigmatism of oblique pencils in addition.

In the experiment of fig. 10 (Plate IX) the object G consists of concentric circles and intersecting radii ruled on a ground glass disc, which is illuminated from behind. At a short distance from the object is situated the strongly curved convex lens L, which produces a real image of the object. This image is received on a lightly powdered disc of plate glass S_1, which is adjusted so that the centre of the circles and the innermost circle appear sharp. With this position of the screen the images of the remaining circles are blurred. If, however, the screen is moved nearer the lens, so that it occupies the position S_2, the images of the outer circles become sharp, whereas the image of the parts of the object near the axis become blurred. In making the figure, two plate-glass plates were used simultaneously and a photograph was taken of the whole, so that the two image planes are shown together in the figure.

From the rear plate S_2 we see that the images of a circle and the part of a radius cutting it cannot be sharp at the same time, the sharp portion of the radius lying nearer the edge than the sharply-defined circle. This difference is due to the astigmatism of the rays, which pass very obliquely through the lens (see below). The same thing is shown very clearly in the similar photograph in fig. 11 (Plate IX).

The blurring of the images in the experiments illustrated by figs. 10, 11 is due to the fact that the pencils of rays forming the image are not reunited in one point on the screen, the position of closest approach lying behind or in front of the screen, so that the section intercepted by the screen is always that of a pencil which is either still converging or has already begun to diverge. If there were no astigmatism, this section would be a simple circle (circle of confusion). As a result of astigmatism, however, a peculiarly shaped spot of light (coma †) is formed. If the angle of the cone of rays forming the image is made very small, the circle of confusion likewise becomes small, and hence the figure on the screen becomes more and more like a sharp image the smaller the aperture of the cone of rays.

* The lens is hidden by its mount and so does not appear in the photograph. In figs. 12–14 the lens and stop have been added to the figure subsequently.

† Lat., *coma*, hair.

In Chap. I, § 4 (p. 9; figs. 9, 10) we produced a figure, comparable to an inverted image, by means of a pin-hole camera and a luminous object, no lens being used. This figure arose from the fact that the minute hole selected a very narrow pencil of rays from the totality of rays emitted by the points of the luminous object. It follows that the "image" may be rendered sharp over a larger region by interposing a stop somewhere in the path of the rays. The use of a stop enables even the blurred images of objects of considerable depth, i.e. with points at varying distances from the lens, to be made sharp.

The effect of a stop on a complex of rays produced by a strongly curved convex lens may be investigated by inserting the stop in the path of the rays in the experiment of fig. 12 (Plate X), which in its essentials resembles that of fig. 9. In the experiment of fig. 13 (Plate X) the stop B is set up close to the lens on the *image* side. We see that as a result the pencils of rays are more sharply bounded and that an image consisting of seven fairly sharply-defined spots of light appears on the powdered glass plate on which the image is again received. We can see quite clearly that the rays producing the image were already present in fig. 12, and that the greater clearness in fig. 13 arises from the fact that the stop lets through only the *outermost* portions of the individual cones of rays.

If, as in fig. 14 (Plate X), we put a stop on the *object* side of the lens, a satisfactory image is again obtained, the stop letting through only the *innermost* parts of the cones of the rays which without the stop would pass through the whole of the lens.

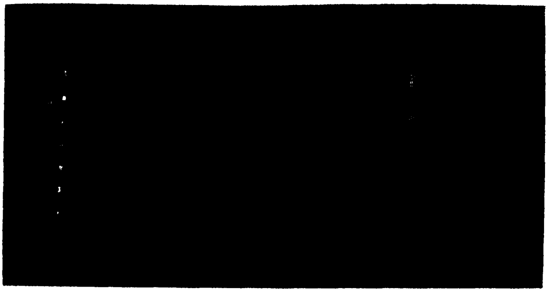
If we compare figs. 13 and 14 we see that the images produced on the same screen are not the same, that in fig. 13 being more spread out and that in fig. 14 more compressed towards the centre. We must emphasize once again, however, that the stop is not the part of the apparatus which actually produces the image, but that the lens of the experiment shown in fig. 12 produced the two images in figs. 13 and 14, the stop merely serving to blot out parts of the figure already existing.

This peculiar action of a stop is shown in a particularly striking way if the object consists of a system of two sets of parallel lines crossing at right angles.

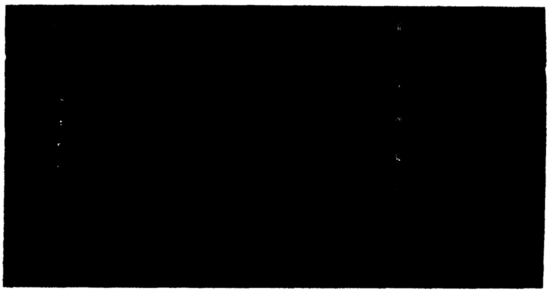
Figs. 15, 16, and 17 (Plate XI) show the images formed of this network by a sharply-curved lens. Fig. 15 was obtained without the use of a stop. In fig. 16 a stop was placed on the image side of the lens, in fig. 17 on the object side of the lens. We see that in fig. 15 the whole image is blurred, but particularly the outer parts.

The indistinctness is diminished by the interposition of a stop; but there results the peculiar distortion of fig. 16, in which the marginal parts are spread out (**cushion-shaped distortion**), while if the stop is placed on the object side of the lens the marginal parts of the image are compressed towards the centre (**barrel-shaped distortion**). The explanation of these types of distortion is contained in the experiments

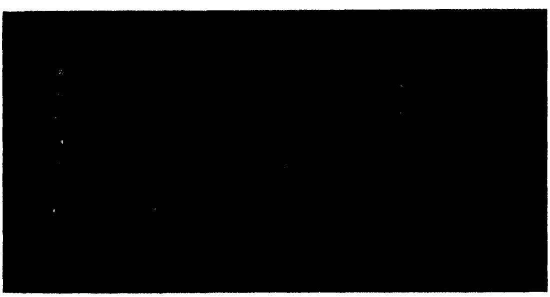
PLATE X



Ch. V, Fig. 12. Formation of an image in the absence of a stop



Ch. V, Fig. 13.—Formation of an image with a stop on the image side of the lens



Ch. V, Fig. 14.—Formation of an image with a stop on the object side of the lens

of figs. 13 and 14, but will be made clearer by the diagrams in figs. 18 and 19.

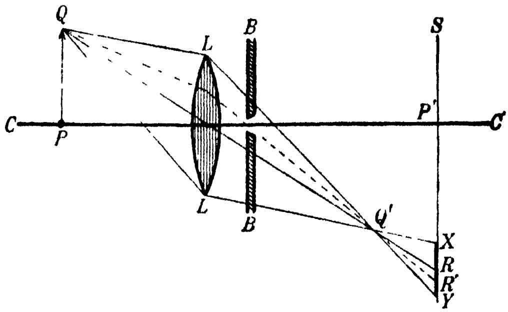


Fig. 18.—Effect of a stop on the image side of the lens

In fig. 18 CC′ denotes the axis of the lens LL, which is to produce an image of the object PQ at right angles to the axis. We set up the screen S in such a way that P′, the image of the point P on the axis, is sharp. As a result of aberration and astigmatism the rays starting from the object-point Q have their point of closest subsequent approach at Q′, which lies nearer the lens than the screen does. Beyond Q′ the rays diverge again and give rise to a bright spot of light, indicated in the figure by XY. The centre of this spot is R; this point is produced by the central ray, which passes through the centre of the lens. If we now insert the stop BB on the image side of the lens a narrow cone is selected from the whole pencil of rays starting from Q; this now gives rise to a much smaller spot of light R′ on the screen S. If R′ coincided with R, R′ would if sufficiently small produce an image which would be geometrically similar to the object; but R′ lies farther from the centre of the screen than R; hence the resulting image has its marginal parts spread out, i.e. the image is distorted.

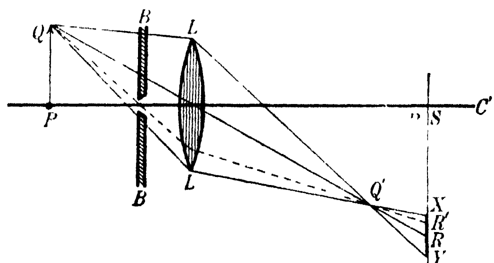


Fig. 19.—Effect of a stop on the object side of the lens

The degree of distortion is greater the greater the distance of the object from the axis of the lens, since the point of closest subsequent approach of rays from points remote from the axis is nearer the lens than that of rays from points near the axis.

The explanation of the barrel type of distortion resulting from the interposition of a stop on the object side of the lens is sufficiently clear from fig. 19. In this figure all the letters correspond to those in fig. 18. We easily see that a stop on the object side lets through only those parts of the pencil of rays that produce a spot of light on the screen lying nearer the axis of the lens than the point R which would give a geometrically similar image.

It follows from these considerations that a stop cannot fail to

cause distortion unless it is inserted in the path of the rays in such a way as to let through only those rays which pass through the optical centre of the lens. This is only possible in the case of optical systems containing at least two lenses; the stop must then be placed between the two.

Here and above it has, it is true, been assumed that the principal points of the system are free from spherical distortion. As, however, photographic lenses of symmetrical form, in which the central rays go through the principal points, give rise to very considerable distortion, the above conclusions must be modified according to the degree of distortion.

The Tangent Condition.—An image free from distortion of any kind (an *orthoscopic* image) must have the same magnification at every point. If PRQ is the object, P′R′Q′ the corresponding image (fig. 20), we must therefore have

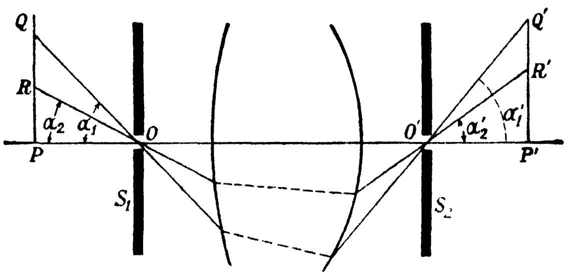
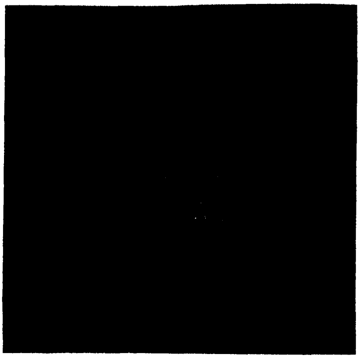


Fig. 20.—Formation of an undistorted image of O at O′ by an optical system with two stops requires the tangent condition $\tan\alpha_1/\tan\alpha_2 = \tan\alpha_1'/\tan\alpha_2'$

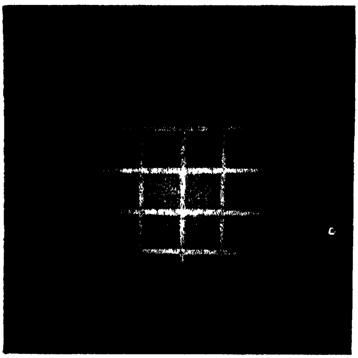
$P'R'/PR = P'Q'/PQ$ or $PQ/PR = P'Q'/P'R'$. Now suppose there is a stop S_1 between the object-point P on the axis and the optical system, and let the point O on the axis be at the centre of its opening. Similarly, let there be a corresponding stop S_2 at the point O′ on the axis corresponding to O, between the optical system and the point P′ corresponding to P. Then the central rays of the pencils starting from Q and R must intersect at O before refraction and at O′ after refraction. If the central rays QO and RO are inclined at angles α_1 and α_2 to the axis and the corresponding central rays after refraction at the angles α_1' and α_2', we have $\tan\alpha_1 = PQ/PO$, $\tan\alpha_2 = PR/PO$, and $PQ/PR = \tan\alpha_1/\tan\alpha_2$. The central rays determine the points Q′, R′ on a screen perpendicular to the axis at P′, and hence, by the foregoing, determine P′R′Q′, the image of PRQ. Hence we must also have $P'Q'/P'R' = \tan\alpha_1'/\tan\alpha_2'$. The *condition for distortionless reproduction* which we gave above accordingly takes the form:

$$\frac{\tan\alpha_1}{\tan\alpha_2} = \frac{\tan\alpha_1'}{\tan\alpha_2'}.$$

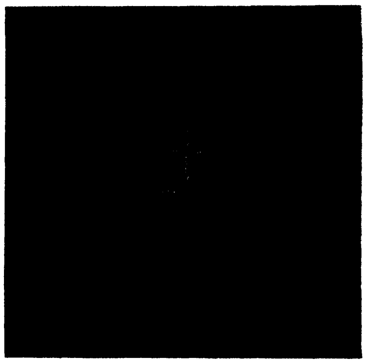
In technical optics this condition, which has to be satisfied by rays inclined at a large angle to the axis, is called the **tangent condition**. It must not be confused with Helmholtz's equation (p. 77); the tangent condition is a statement about the inclinations of the central rays at *corresponding points* on the axis, O and O′, whereas Helmholtz's equation is a statement about the inclinations of *corresponding rays* at the points P and P′ on the axis.



Ch. V, Fig. 15 - Indistinct image of a network obtained without a stop



Ch. V, Fig. 16 - Image of a network obtained by using a stop on the image side of the lens (cushion-shaped distortion)



Ch. V, Fig. 17 - Image of a network obtained by using a stop on the object side of the lens (barrel-shaped distortion)

Ch. VI, Fig. 13. - Foci in the Galilean telescope (p. 126)

5. Refraction Phenomena considered from the point of view of the Wave Theory of Light.

Light Ray and Wave Normal.—All our previous discussion has been based on the assumption that we can regard the path along which light energy travels as a geometrical straight line. It is, however, very instructive to consider the phenomena of refraction on the basis of the wave theory of light, and this point of view is very important if we are to gain a deeper understanding of refraction phenomena and a clear recognition of the limitations of the methods of geometrical optics. We are enabled to pass from one point of view to the other by the fact that the light ray, on which our constructions have hitherto been based, at every point of its course represents the normal to the wave front.

The actual way in which refraction is produced when wave motion crosses the boundary of a medium has already been dealt with in Vol. II. (p. 248). In what follows we shall consider the most important special cases which have a special bearing on optics.

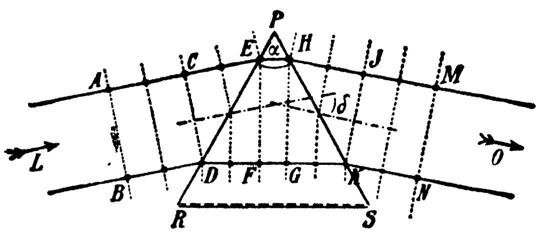


Fig. 21.—Refraction of parallel light by a prism

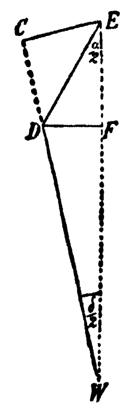


Fig. 22

Refraction of Light Waves by a Prism.—We consider the case where the path of the ray is symmetrical (fig. 21). The wave surface of the parallel pencil of rays at a definite instant is the plane AB. It is moving forwards. In the position CD the lower part of the wave surface has already reached the prism PRS and thenceforward advances with diminished velocity. The wave surface (wave front) is gradually turned into the position EF and thereafter advances parallel to itself until its upper part again leaves the prism at H and travels on with increased velocity in air. There is a second rotation of the wave until the whole pencil of rays is advancing in the new direction from the position JK.

To calculate the quantitative relationship between α, the refracting angle of the prism, δ, the minimum deviation of the ray, and μ, the refractive index, we redraw part of the figure separately in fig. 22. CE is the direction of the incident ray, DF the direction of the ray in the prism, i.e. the angle formed by CE

and DF is half the deviation of the pencil. The same angle $\delta/2$ is formed by the wave fronts CD, EF, which are at right angles to the directions of the rays and whose continuations meet at W. Thus

$$\angle DEF = \frac{\alpha}{2}, \quad \angle DWE = \frac{\delta}{2},$$

so that

$$\angle CDE = \frac{\delta + \alpha}{2}.$$

Then

$$\sin CDE = \sin \frac{\delta + \alpha}{2} = \frac{CE}{DE},$$

$$\sin DEF = \sin \frac{\alpha}{2} = \frac{DF}{DE}.$$

By division $\sin \frac{\delta + \alpha}{2} \Big/ \sin \frac{\alpha}{2} = CE/DF$. The ratio CE/DF is the ratio of the velocities of propagation of light in air and in the prism, for the paths CE and DF are described by the light in the same time. Now this latter ratio is equal to the refractive index of the substance of the prism; hence

$$\frac{\sin \dfrac{\delta + \alpha}{2}}{\sin \dfrac{\alpha}{2}} = \mu,$$

in agreement with the equation deduced in Chap. III, § 6 (p. 62).

The Optical Path.—Another noteworthy principle may [illegible] derived from fig. 21. Along the pencil of rays LO a num[illegible] [illegible]ve fronts have been drawn, the distance between each [illegible] [illegible]ccessive fronts being so chosen that the time taken between [illegible]o successive positions is always the same. We see at once from the figure that the ray of light following the path ACEHJM requires the same time to traverse the distances between the wave fronts AB and MN as the ray BDFGKN does. Any other ray belonging to the pencil requires the same time. It follows that in order to fix the path followed by a pencil of rays in any medium, whether homogeneous or compound, we have the following principle:

All parts of a pencil of rays take the same time to traverse the medium between one wave front and another.

If we bear in mind that a wave front is the geometrical locus of all points vibrating in the same phase, there can of course be only one definite difference of phase between the points of two wave fronts of a pencil of rays. This phase difference, moreover, is a measure of the time taken by the wave front to move from one position to the other. Thus the principle stated above is immediately obvious.

This principle is particularly fruitful when we are tracing the paths of the rays of a homocentric pencil of rays (p. 80) through various media by considering the paths of individual rays from one centre to the other. For at the

centres the wave surfaces of the homocentric pencils shrink to points. Hence the centres in which all the rays of the pencil intersect without aberration may themselves be regarded as wave surfaces. Hence the time taken for light to move from one centre to the other along any ray joining the two centres must be the same. If a light ray moving with velocity c_1 in the first medium describes a path of length l_1, the time required is l_1/c_1. If the length of path in the second medium is l_2 and the velocity of light c_2, the time taken in this medium is l_2/c_2. Hence to pass from one centre to the other the ray of light in question requires the time $\frac{l_1}{c_1} + \frac{l_2}{c_2} + \ldots + \frac{l_m}{c_m}$. Suppose that another ray of the same homocentric pencil traverses the distances $l_1', l_2', \ldots, l_m'$ in the same media. Then according to our principle we must have

$$\frac{l_1}{c_1} + \frac{l_2}{c_2} + \ldots + \frac{l_m}{c_m} = \frac{l_1'}{c_1} + \frac{l_2'}{c_2} + \ldots + \frac{l_m'}{c_m}.$$

If we multiply the equation by c, the velocity of light in a vacuum, and remember

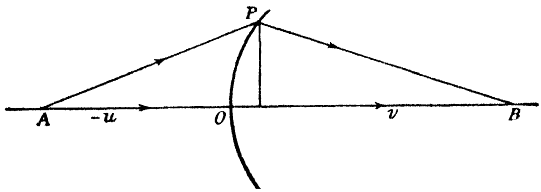


Fig. 23.—Formation of an image without aberration

that $c/c_1 = \mu_1$, $c/c_2 = \mu_2$, &c., where μ_1 is the (absolute) refractive index of the first medium, &c., we obtain

$$l_1\mu_1 + l_2\mu_2 + \ldots + l_m\mu_m = l_1'\mu_1 + l_2'\mu_2 + \ldots + l_m'\mu_m;$$

that is, for all rays of the homocentric pencil

$$\sum_i l_i\mu_i = \text{const.}$$

Conversely, this condition is the sole condition for a homocentric pencil, i.e. a pencil such that all its rays pass through one point *without aberration*. If we call the product of the refractive index of a medium and the length of the path of the light in the medium the *optical path* (or *reduced path*), we may state the condition in words as follows:

One point cannot be the aberrationless image of another (i.e. the rays from one point cannot be accurately recombined at the other) unless all the possible optical paths between the two points are equal.

Formation of an Aberrationless Image of a Point by a Refracting Surface.— The theorem just stated makes it comparatively easy to discuss a problem which we have already referred to repeatedly, namely, the determination of a surface which will form an image at a given point B of a given object A, in which the rays are accurately recombined. Obviously this surface will be a surface of revolution with AB as axis; let it meet the axis at the point O (fig. 23). Let μ

be the refractive index of the medium surrounding the object and μ_1 that of the medium surrounding the image. Then all the rays starting from A will not meet accurately at B unless the sum of the optical paths, $\mu AP + \mu_1 PB$, is independent of the position of the point P where the ray AP starting from A meets the refracting surface. Hence we must have

$$\mu OA + \mu_1 OB = \mu PA + \mu_1 PB$$

no matter what the position of P is. We denote OA by u and OB by v, and starting from O we reckon these lengths positive in the direction in which the light is moving (p. 40). Hence in fig. 23 we take OA as $-u$. We now take O as the origin of a system of rectangular axes and the optical axis AB as x axis; let the co-ordinates of P be x, y. Then $PA = -\sqrt{\{(-u + x)^2 + y^2\}}$, $PB = \sqrt{\{(v - x)^2 + y^2\}}$, and we obtain the fundamental equation

$$-\mu u + \mu_1 v = -\mu\sqrt{\{(u - x)^2 + y^2\}} + \mu_1\sqrt{\{(v - x)^2 + y^2\}}$$

as the condition for freedom from aberration. If we interpret x, y as current co-ordinates, the equation represents a curve of the fourth degree.* Thus we have the following result:

An aberrationless point image of a point can be produced by a surface of revolution whose cross-section is a curve of the fourth degree.

The fundamental equation of the cross-section includes all particular cases of the formation of an aberrationless image.

(1) We first put $\mu_1 = -\mu$; instead of refraction we then have the particular case of regular reflection. At the same time we must replace v by $-v$ (see above). Hence a mirror which reflects without aberration has a cross-section given by the equation

$$u - v = \sqrt{\{(u - x)^2 + y^2\}} + \sqrt{\{(-v - x)^2 + y^2\}}.$$

This may be transformed into

$$\left(x - \frac{u - v}{2}\right)^2 - \frac{y^2(u - v)^2}{4uv} = \left(\frac{u - v}{2}\right)^2.$$

According as uv is positive or negative, this is the equation of a hyperbola or an ellipse, the foci being A and B. If we let u increase without limit, i.e. let A move to infinity, the curve becomes a parabola. If $u - v = 0$, i.e. if O is the mid-point of AB, the equation becomes

$$x = 0,$$

i.e. the curve becomes the perpendicular to AB at O. Thus the ellipsoid of revolution is the only reflecting surface which will form a real image of a real object without aberration, and the hyperboloid of revolution the only reflecting surface which will form a virtual image of a real object without aberration. The paraboloid of revolution represents a limiting case of both, and the plane mirror is a particular case of the paraboloid of revolution.

(2) We now make the left-hand side of the fundamental equation equal to zero, i.e. we take the point O between A and B in such a way that it divides AB in the inverse ratio of the refractive indices. The equation then becomes

$$0 = -\mu\sqrt{\{(u - x)^2 + y^2\}} + \mu_1\sqrt{\{(v - x)^2 + y^2\}}.$$

* The curves are known as Cartesian ovals, because the problem in question was discussed and solved by DESCARTES.

If we put $\mu_1 v = \mu u$, this may be transformed into

$$\left(x - \frac{\mu u}{\mu + \mu_1}\right)^2 + y^2 = \left(\frac{\mu u}{\mu + \mu_1}\right)^2.$$

This is the equation of a circle. Hence the surface which will refract without aberration is a *sphere*. We thus obtain the important optical theorem:

For every spherical refracting surface there are two points such that one is the aberrationless image of the other

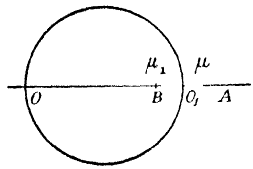


Fig 24 —The aplanatic points of a spherical surface

The radius of the sphere is $\mu u/\mu + \mu'$. The points A, B, O, and O_1 (the second point of intersection of the spherical surface with the axis (fig 24)) accordingly determine the distances

$$OA = u, \quad OB = v = \frac{\mu u}{\mu_1}, \quad OO_1 = \frac{2\mu u}{\mu + \mu_1},$$

$$O_1B = \frac{\mu u}{\mu_1} - \frac{2\mu u}{\mu + \mu_1} = -\frac{u(\mu_1 - \mu)\mu}{\mu_1(\mu_1 + \mu)},$$

$$O_1A = u - \frac{2\mu u}{\mu + \mu_1} = \frac{u(\mu_1 - \mu)}{\mu_1 + \mu}.$$

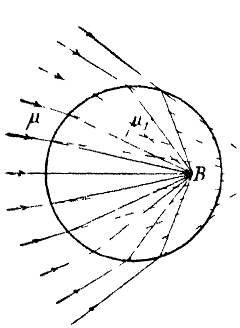


Fig 25

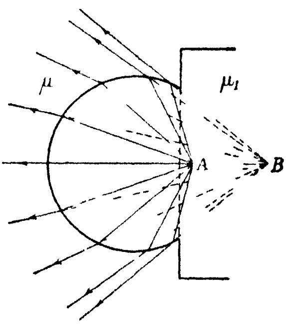


Fig 26

The two cases of aplanatic refraction by a spherical refracting surface

These lengths have the ratios

$$\frac{OA}{O_1A} = \frac{u(\mu_1 + \mu)}{u(\mu_1 - \mu)} = \frac{\mu_1 + \mu}{\mu_1 - \mu},$$

and

$$\frac{OB}{O_1B} = -\frac{v\mu\mu_1(\mu_1 + \mu)}{\mu_1 u(\mu_1 - \mu)\mu} = -\frac{\mu_1 + \mu}{\mu_1 - \mu}.$$

In geometrical language, therefore, the points A, B, O, O_1 form a harmonic range. Hence for a given spherical surface they may be found by dividing the diameter of the sphere in the ratio of the sum and difference of the refractive indices. In figs. 25 and 26 the positions of A, B are given for $\mu = 1$, $\mu_1 = 1{\cdot}5$. These two special points are called the **aplanatic points** (or **aplanatic foci**) of the sphere. As we see from the two figures, either the object-point is virtual and the image-point is real, or, as object and image are always interchangeable,

the object-point may be real and the image-point virtual. In order that such a path should be actually realizable when the refractive index of the sphere is the greater, i.e. when μ_1 is greater than μ, it is necessary that only a part of the spherical surface should be used; this is done by embedding the object A in a medium with the same refractive index as that of the sphere (fig. 26). This case is exemplified in the front lens of the microscope apochromat objective (fig. 28, p. 140) due to E. ABBE.

(3) In all *other* cases it is *impossible* to obtain wholly aberrationless refraction with a spherical surface. If, however, we further limit the case to rays which meet the axis in the neighbourhood of O, we may regard y^2 in the fundamental equation as small compared with $(u-x)^2$ and $(v-x)^2$. If we take into account only terms of the second degree in $y/(u-x)$ and $y/(v-x)$, the expansion of the fundamental equation

$$-\mu u+\mu_1 v=-\mu(u-x)\sqrt{\left\{1+\left(\frac{y}{u-x}\right)^2\right\}}+\mu_1(v-x)\sqrt{\left\{1+\left(\frac{y}{v-x}\right)^2\right\}};$$

gives approximately

$$0\approx\mu x-\mu_1 x-\frac{\mu y^2}{2(u-x)}+\frac{\mu_1 y^2}{2(v-x)}.$$

Dividing the equation by $y^2/2$ and rearranging, we have

$$\frac{\mu_1}{v-x}\approx\frac{2x(\mu_1-\mu)}{y^2}+\frac{\mu}{u-x}.$$

If we now put $y^2=2rx$, i.e. if we make the curvature of the refracting surface parabolic, so that the parabola has latus rectum $2r$, i.e. radius of curvature r at the vertex, the equation takes the form

$$\frac{\mu_1}{v-x}\approx\frac{\mu}{u-x}+\frac{\mu_1-\mu}{r}.$$

This equation becomes identical with the object-image relationship

$$\frac{\mu_1}{v}\approx\frac{\mu}{u}+\frac{\mu_1-\mu}{r},$$

or $$V=U+D,$$

if we may also neglect x as compared with u and v (when r is large).

From these considerations it follows that a parabolic surface—and likewise a spherical surface, provided only that such small portions of the surface are used that they are indistinguishable from a parabolic surface (r large) – represents an approximation to a surface giving aberrationless refraction, provided there is a limitation to axial rays. But whereas the general surface of the fourth degree is only devoid of aberration for a *single* pair of points on the axis, these approximation surfaces are capable of forming aberrationless images of a considerable part of the axis, namely, so long as the object-point and the image-point do not approach so close to the surface that x no longer vanishes compared with u and v. Thus we have on the one hand established our former object-image theorems on the basis of the optical path and the ideas of physical optics, and on the other hand put a definite limitation to their range of validity.

Refraction of Light Waves at Lenses.—Let the point source G lying on the axis of a convergent lens (fig. 27) emit a homogeneous pencil of rays. If we may assume that B is an aberrationless image of G, the pencil of rays starting from G must on refraction at the lens

become the homocentric pencil of rays converging to the point B. The ray GN_1N_2B along the axis takes the same time to travel from

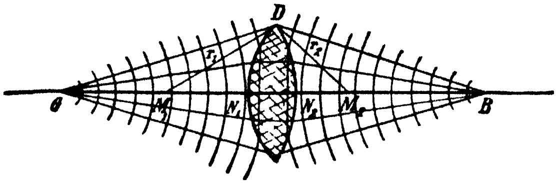


Fig. 27.—Formation of an image by a lens; wave fronts for a number of definite instants

G to B as the ray GDB. The positions of the wave-fronts are shown in the figure for a number of definite instants.

Curved Light Rays.—If the refractive indices differ only very slightly from one another, the change in direction of the light ray is also small. If a ray of light traverses a number of bodies in succession, such that each one has a refractive index differing by only an infinitesimal amount from that of the preceding, i.e. if the refractive index varies continuously, the change in direction of the light ray is also continuous, so that the ray follows a curved path.

We can convince ourselves of the truth of this by a simple experiment. Into a large plate-glass dish with a flat bottom we pour water and then cautiously add a layer of alcohol, so that there is a clearly visible plane boundary between the two liquids. If a ray of light falls obliquely on the boundary surface, it is refracted in such a way that its angle of incidence on the alcohol is less than that in the water, as the refractive index of alcohol is greater than that of water. If we now let the dish stand undisturbed for a considerable time, the liquids gradually diffuse into one another (the mixing may also be brought about by careful stirring), the sharply-defined boundary vanishes, and an extensive region of transition is formed, in which each successive layer has a refractive index slightly exceeding that of the layer below it. The ray of light in passing from the uppermost layer consisting of almost pure alcohol to the lowest layer consisting of almost pure water is now subject to a gradual curvature. If we then let the ray of light fall on the liquid in such a way that it makes a very small angle with the successive layers, the curvature of the ray may be such that it turns upwards again (fig. 28). If the alcohol is coloured by means of a little fluorescein, the path of the light is marked out by bright green fluorescence and may thus be observed very clearly.

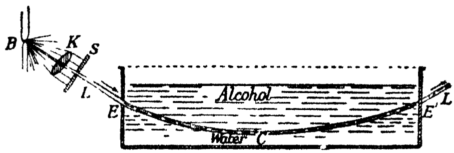


Fig. 28.—Curved path of a ray of light passing through a layer of alcohol on top of water

The explanation of the curvature of the path of the light on the wave theory readily follows from fig. 29. The pencil of parallel rays L incident from the left has at a definite instant the plane wave-front A_1A_2. Now as the refractive index is greater in the upper layers of the liquid than in the lower ($\mu_1 > \mu_2$), the velocity of the light in the upper part of the pencil, i.e. at A_1, is less than in the lower part, at A_2. Hence the wave-front at a place B_1B_2 reached by the light ray later

takes a form determined by the fact that A_1B_1 is less than A_2B_2. It follows in the same way that the wave-fronts at the points reached subsequently are inclined to the wave-fronts at points reached previously. Now as the directions of the rays of light are always at right angles to the wave-fronts, it follows that the pencil follows a curved path. The pencil leaves the liquid in the direction L′. If we produce the directions of the incident and refracted rays, they intersect at R. The pencil of rays accordingly leaves the liquid as if it were reflected at R. To an eye looking obliquely into the liquid the layer SS appears like a totally reflecting plane mirror.

If (as in fig. 30) a number of rays start from a luminous point P in a stratified liquid, all those rays which meet the layers at only a small angle travel as shown

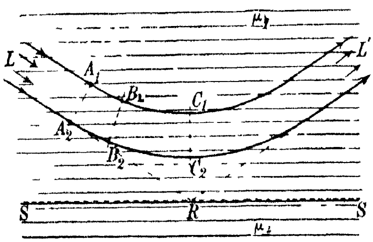


Fig. 29.—" Reflection " due to curvature of the light rays

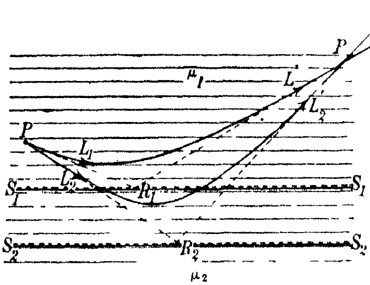


Fig. 30 —Double image due to curvature of the light rays

in fig. 28. Now it may happen that two rays L_1 and L_2 starting from P are curved in such a way that they meet again at a point P′. If the eye of an observer is placed at P′, the two rays enter the eye in different directions (L_1' and L_2'), and hence the eye sees images of the luminous point in the backward prolongations of both rays. These images are situated as if the rays R_1 and R_2 starting from P were reflected at the apparent reflecting surfaces S_1S_1 and S_2S_2.

We shall come back to this phenomenon when explaining mirage (Chap. XI, § 1, p. 262).

Effect of Diffraction.—The occurrence of refraction phenomena as a result of the velocity of the wave motion of light, which we have discussed in this section, is insufficient to explain all the relationships involved in the formation of an image. It is only when we take into account the interference of elementary waves which occurs in all these phenomena that we are in a position to give a complete account of the matter (Chap. VI, § 9, p. 144).

CHAPTER VI

Geometrical Optics: Applications

1. Some Applications of Convex Lenses.

The Photographic Camera.—The photographic camera (fig. 1) is a light-tight box; its side walls often consist of a folded leather bag, which enables the distance between the front and back of the camera to be varied. The front wall has an opening into which is fixed a convex lens or a collective system of lenses (acting like a convex lens), called the **objective** (or simply the **lens**); the back wall consists of a **ground-glass screen** (focusing screen), which is replaced by the photographic plate when a photograph is being taken.

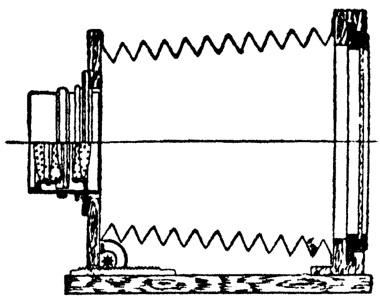


Fig. 1.—A photographic camera

If we set the camera facing an object, the lens if properly focused will form a real image of the object. If the object is very distant the focusing screen must coincide with the focal plane of the lens. If the object is nearer, the camera must be extended, for the distance between the lens and the focusing screen must be increased in accordance with the equation $1/v = 1/u + 1/f$, if a sharp image is to be produced.

Defects in the Image.—In taking a photograph, the image produced must be similar to the object and all parts of the image plane must be sharp. This, however, cannot be attained by means of a single lens, as the latter will always exhibit all the defects discussed in detail in the previous chapter: spherical aberration, astigmatism, and distortion.

In addition, there is **chromatic aberration,** which we shall not be in a position to discuss until we have dealt with the chromatic dispersion of light (p. 166).

The defects of the objective which we have mentioned are all due to the fact that some of the rays forming the image traverse the lenses too near their edges, or that the objective receives rays from points remote from its axis, so that these rays are inclined at too great an angle to be regarded as axial rays. These defects can be partially got rid of by using a convergent *system of lenses* instead of a single lens, and, further, the effect of the marginal rays on the formation of the image may to a certain extent be prevented by inserting a stop or diaphragm.

We explained in Chap. III, § 4 (p. 47) how stops or the mounts of mirrors

and lenses affect the brightness or size of the field of view. According to that discussion, we must first determine the position and magnitude of the aperture stop and then the field of view from the centre of the aperture stop.

If a stop is placed immediately in front of the lens, it acts both as **entrance pupil** and **aperture stop.** If, on the other hand, the stop is inside the lens system, its **aperture** is most simply found in practice by focusing for an object at infinity and then replacing the focusing screen by an opaque screen with a small hole in the centre, lighted from behind. The rays from this point source leave the lens as a parallel pencil, giving rise to a bright circle on a piece of tracing-paper held immediately in front of the lens. This circle may be regarded as the aperture or entrance pupil; let its diameter be d.

The **field of view** depends chiefly on the diameter of those lens mounts which act as field stops. The field of view on the image side may be found experimentally by determining the size of the bright circle produced on the focusing screen when the lens is pointed at the sky.

Any point of the entrance pupil established by experiment may be regarded as the vertex of a cone of rays whose base is the whole field of view on the image side. It follows that the **brightness** of the image on the focusing screen depends in the first place on the number of points lying in the entrance pupil, that is, on the area of the entrance pupil; it is directly proportional to this area, i.e. proportional to the square of d, the diameter of the opening. The brightness of the image is also inversely proportional to the square of the distance of the illuminated surface from the source of light, i.e. in this case the distance between the focusing screen and the lens. For distant objects the distance between the focusing screen and the lens is equal to f, the focal length of the lens; hence the brightness of the image on the focusing screen is proportional to the square of the quotient of the diameter of the opening and the focal length. This quotient $q = d/f$ is called the **relative aperture.** The greater the relative aperture of a system, the brighter the resulting image. Nowadays it is possible to produce satisfactory photographic lenses with a relative aperture of 1 : 1.

Finally, owing to losses of light by reflection and otherwise, the brightness of the image also depends on the nature and number of the lenses of which the objective consists; it is greater the smaller the number of lenses. It is very important to know the relative aperture of a photographic objective, as on this depends the *time of exposure*, i.e. the time which is necessary to illuminate a photographic plate properly. For objectives of any one type we may make the following statement:

The brightness of the image is proportional to the square of the relative aperture. Hence E, the time of exposure, is inversely proportional to the square of the relative aperture:

$$E = \frac{k}{(d/f)^2},$$

where k is a constant of the lens system.

The Magnitude of the Image.—Let an object of magnitude y be at a great distance u from the first principal plane. Let the ray from the tip of the object to the first principal point form the angle φ with the axis. Then $\tan \varphi = y/u$. As the medium is the same (air) on both sides of the lens, the principal points and the nodal points coincide. The prolongation of the ray beyond the second principal point is therefore inclined at the same angle to the axis. As the distance u is taken very large, the image of magnitude y' is situated at the focus. Hence $\tan \varphi = -y'/f$, so that $y' = -f \tan \varphi = -fy/u$.

Here φ is the *apparent magnitude* of the object (p. 79). For example, the apparent magnitude of the sun is 31′. Hence in a photograph of the sun the diameter of the image is $y' = f \tan 31'$. This could be used to determine the

focal length f. It is clear that to obtain large images systems of large focal length must be used.

The Projection Lantern.—While the photographic camera forms a real image of a distant external object on the focusing screen or photographic plate in the neighbourhood of the second focus, the projection lantern * forms a real (magnified) image of an object (lantern slide) placed within the apparatus near the focus, on a screen outside the apparatus.

As with any lens the positions of the object and image are interchangeable, any photographic camera may be used as a projection lantern if the focusing screen is replaced by a lantern slide and the screen is set up at a suitable distance from the camera.

In projection on a magnified scale the quantity of light emitted by the object is spread over a larger area. Hence if the image on the screen is to be sufficiently bright, the slide must be illuminated very brightly. A particularly powerful source of light is therefore used and the rays are directed in such a way that as much of the light as possible is utilized in producing the image.

The projection apparatus is accordingly arranged as follows (fig. 2). A powerful source of light L is placed in a light-tight box. In an opening in the box is fixed the **condenser** C, which is frequently a combination of two plano-convex lenses. At a definite distance from the condenser is fixed the objective O with the stop Bl inside. The cone of rays indicated by dots in the figure is made to converge by the condenser, so that a real image of the source of light is formed at the projection lens. The slide GG′ is placed immediately in front of the condenser; it is illuminated very brightly by the cone of rays from behind. A real image of the slide GG′, which acts as object, is formed by the objective on the screen SS′. The central ray of the cone of light which starts from the point A of the object and gives rise to the point B of the image is drawn in the figure.

In most cases the screen SS′ is so far away that the object GG′ must be placed approximately at the first focal plane of the objective: hence the distance of the objective from the front of the condenser is approximately equal to the focal length of the objective. If all the

* The projection lantern is a development of the "magic lantern". The German Jesuit Athanasius Kircher was the first to set up various types of apparatus (1646 onwards) by which magnified images could be thrown on a wall by means of a concave mirror (acting as a condenser) and a projection lens, using sun- or candle-light. The images to be projected were actually visible in the mirror in this case. About 1653 the mathematician Andreas Tacquet displayed "Martinus' journey from China to the Netherlands" at Löwen by means of Kircher's apparatus. Probably this was the first lantern lecture ever delivered. The first-known magic lantern with interchangeable pictures painted on flat glass plates (slides) and a two-lens objective was demonstrated by Thomas Walgenstein, a Danish student at Leyden who probably got the idea from Tacquet, in a number of cities in 1658 and subsequently, and created a great sensation. C. Huygens, who had dealings with Walgenstein, also possessed a proper magic lantern about that time. He even used a lens as condenser, whereas Walgenstein still retained the use of a concave mirror.

rays which are made to converge by the condenser C are utilized in producing the image, the cone of rays must converge to the entrance pupil of the objective; hence it is not a matter of indifference what the focal length of the condenser is. In order to make the rays travel

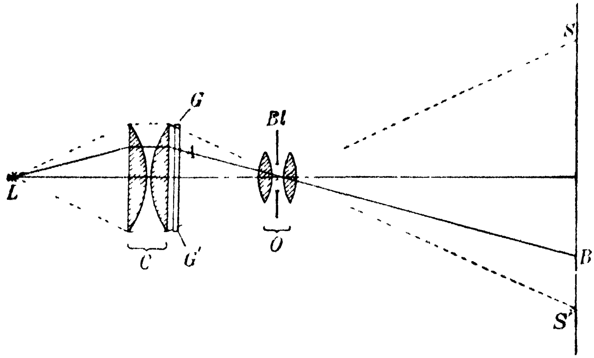


Fig. 2.—Path of the rays in the projection lantern

symmetrically through the condenser, we bring the point of convergence of the rays passing through the condenser as far in front of the condenser as the source of light is behind it, i.e. the position is chosen so that the point of convergence of light passing through the condenser is at a distance from the condenser equal to twice the focal length of the condenser. It follows that the most favourable arrangement is that in which twice the focal length of the condenser is approximately equal to the focal length of the objective.

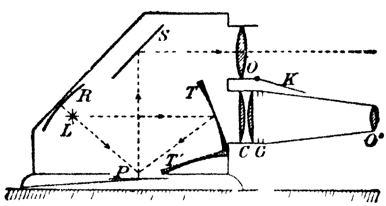


Fig. 3.—An epidiascope

In the so-called **episcope** (see fig. 3), the image is not produced by the light passing through the image; on the contrary, the object P is very strongly illuminated by a powerful source L by means of the concave mirrors R and T, and the light reflected by it is caused to form an image by the very accurate plane metal mirror S and the objective O. The image produced by an episcope is in general much fainter than that produced by the ordinary projection lantern, as the illuminated object absorbs the greater part of the incident rays and scatters them in directions which do not contribute to the formation of the image. Hence episcopes require much more powerful sources of illumination. A combination of the episcope with the projection lantern is known as the **epidiascope** (fig. 3). Here the mirror T is fixed so that it may be rotated into the position T′.

Size of the Image.—We may reverse the steps of the argument used on p. 110 to find the size of the photographic image of a distant object. We now have u as the distance of the image from the second principal plane of the lens, y' the

magnitude of the slide, y the magnitude of the image. Then $y = -uy'/f$ and the magnification $m = y/y' = -u/f$. To obtain large images, therefore, we have to use an optical system of *small* focal length.

2. The Human Eye.

The Structure of the Eye.—The eyeball (fig. 4) consists from the physical point of view of an almost spherical chamber somewhat compressed between back and front, which is provided with a system of refracting substances acting as objective. By means of six muscles it may be rotated in the socket in any direction as if in a ball-and-socket joint.

The outer coat of the eyeball is the white **sclerotic** S, which is very firm and strong and serves to defend the eye from injuries. The front part of the sclerotic is transparent and is known as the **cornea** (C). The sclerotic is coated on the inside with the dark-coloured **choroid** A, which contains both the blood-vessels supplying the eye and a dark layer of pigment which shields the interior of the eye from scattered light. In front the choroid merges into the **iris** I, which has a hole in the centre (the **pupil**). On the inner side the choroid merges into the rose-coloured **retina** R, which consists of the branchings and terminations of the optic nerve. Fig. 5 shows a highly magnified section through the retina. The retina consists of a series of complicated layers. The side indicated by the arrows is next the light. The layer actually sensitive to light is marked L; it consists of a very large number of **rods** R, and **cones** C, which, curiously enough, are directed away from the light.

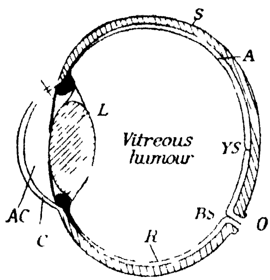


Fig. 4.—A vertical section through the eye

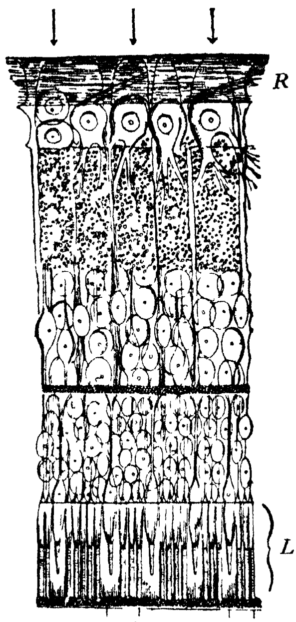


Fig. 5.—A cross-section of the retina

The point of the retina exactly opposite the pupil, known as the **fovea centralis**, contains the greatest number of cones and is the most sensitive portion of the retina. The surrounding region, on account of its colour, is known as the **yellow spot** or macula lutea (YS). At

the point where the optic nerve enters the eye there are no nerve endings. This point (BS) is insensitive to light and is known as the **blind spot.** The blind spot lies on the nose side of the yellow spot.

We can convince ourselves of the existence of the blind spot by looking steadily at the cross in fig. 6 with the right eye, holding the figure about 20 cm. from the eye. The round spot then vanishes, as its image falls on the blind spot (fig. 7). That we are not ordinarily disturbed by the existence of the blind spot is chiefly due to the fact that we look at objects with the two eyes simultaneously.

Fig. 6.—To demonstrate the existence of the blind spot

Behind the iris lies the **crystalline lens** (L in fig. 4), a transparent horny body diminishing in refractive index towards the interior. For the sake of simplicity, we may take 1·4085 as an average value for the refractive index. The crystalline lens has approximately the form of a blunted ellipsoid of revolution, its rear surface being somewhat more strongly curved than the front surface. The crystalline lens divides the interior of the eye into two regions of unequal size; the **anterior chamber** AC between the crystalline lens and the cornea is filled with a colourless liquid (the **aqueous humour**) with a refractive index equal to that of water, while the region between the crystalline lens and the retina contains a transparent jelly-like substance (the **vitreous humour**) whose refractive index is nearly equal to that of water. The line joining the centre of the pupil or the vertex of the cornea to the centre of the yellow spot is termed the **optic axis**; it is at right angles to all the surfaces bounding the refracting parts of the eye.

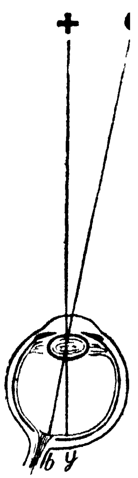


Fig. 7.—To illustrate the experiment with fig. 6

The Eye as an Optical System.—When rays of light enter the eye they are refracted at the three refracting surfaces, the cornea, the front surface of the lens, and the rear surface of the lens. In a normal or **emmetropic** * eye at rest, rays entering parallel to the optic axis are made to converge to the yellow spot; the latter is therefore the focus of the optical refracting system. When at rest the eye sees very distant objects sharply.

Dimensions and Cardinal Points of the Eye.—According to A. Gullstrand, the following are the average values for the emmetropic eye. The length along the axis from the front of the cornea to the yellow spot is about 24 mm. The cornea is 0·5 mm. thick, and its refractive index is 1·376. The radius of curvature

* Gr., *en*, a prefix intensifying the significance of the idea represented by the following syllable of the word; *metron*, a measure; *ops*, the eye, so that *emmetropic* means *seeing accurately*.

of its front surface is 7·7 mm. at the vertex and that of the rear surface 6·8 mm. The liquid in the anterior chamber, the aqueous humour, has the refractive index 1·336. The refracting power of the system consisting of the cornea and the aqueous humour is 43·05 dioptres. The radius of curvature of the front surface of the lens is 10 mm., that of the rear surface 6 mm. From the age of twenty years onwards the lens may be subdivided into an outer or *cortical layer* and a *nucleus*. The mean refractive index of the outer layer is 1·386, that of the nucleus is 1·406. The refracting power of the lens is 19·11 dioptres. The optical system consisting of the cornea, the aqueous humour, and the lens has a total refracting power of 58·64 dioptres. The first principal point lies 1·35 mm. and the second principal point 1·60 mm. behind the front of the cornea. The aforesaid refracting power corresponds to a focal length of 17·06 mm., so that the first focus lies 17·06 — 1·35 = 15·71 mm. in front of the foremost part of the cornea. The vitreous humour behind the lens has the refractive index 1·336. The second focus of the optical system just mentioned lies 24·39 mm. behind the front of the cornea, so that the second focal length is 24·39 — 1·60 = 22·79 mm. (This focal length is 1·336 times the first one; see p. 76.)

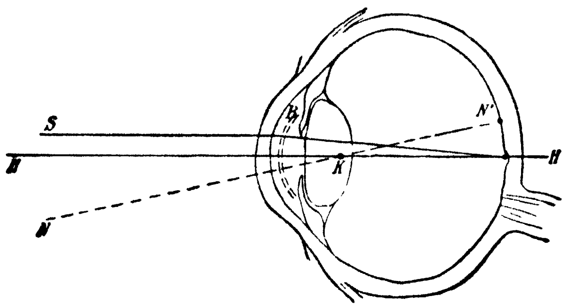


Fig. 8.—Listing's reduced eye (B) and the nodal point K

According to LISTING, the three refracting surfaces may for optical purposes be replaced by a spherical surface B (fig. 8) whose centre of curvature K lies near the rear surface of the lens. Hence the path of a ray of light S entering the eye parallel to the optic axis HH may be constructed by making the ray run parallel to the axis until it meets the spherical surface B and then joining the point of intersection to the yellow spot. The eye thus simplified is called the **reduced eye.** All the rays of light NN′ which pass through the centre of curvature K of the spherical surface just described traverse the eye without change of direction. The point K is called the *nodal point* of the reduced eye. (The actual nodal points of the eye are determined by the focal distances and the positions of the principal planes; cf. p. 83).

The Sensitiveness of the Eye.—The photosensitive retina of the eye serves as a screen for the images produced by the optical system. Not all parts of the retina are equally capable either of detecting minute light stimuli or of recognizing the finer details of an object under examination. The maximum capacities for the two types of sensation are associated with different parts of the retina.

The capacity for recognizing as many details as possible in an object under examination is in other words the capacity of being able to

perceive as separate two small but optically distinct point-objects which are close together. The measure of this power is called **visual acuity.** It may be expressed quantitatively by the reciprocal of the angle included by the rays from two points when these can still be just seen as two (p. 28). Visual acuity is greatest by far at the yellow spot, decreasing rapidly towards the retinal margin. By movements of the eye and head the eye is accordingly directed in such a way that the image of the object to which attention is directed and on which the eye is focused falls on the yellow spot. Vision by means of the yellow spot is called *direct vision*, vision by means of other parts of the retina *indirect vision* (see § 4, p. 121).

If we wish to examine an extended region of the object, we usually cause the eye to rotate rapidly in its socket. The optic axis then successively passes over the various parts of the object under examination, and it is only the succession of accurately-seen partial images that gives us a general idea of the object. The central rays of the pencils entering the eye from separate elements of the object intersect at the centre of the pupil. This, then, is the meeting-point of the visual rays in *indirect* vision, and hence the centre of perspective of the configuration formed by the visual rays. If the eye is *moving*, the central rays for successive positions of the eye meet at the point about which the eye is rotating. These rays may be called the lines of sight.* The centre of the perspective in space formed by the lines of sight, on which our perception of space depends, is the point about which the eye is rotating.

By the **sensitiveness** of the eye we merely mean its capacity to be so stimulated by a definite luminous flux that the stimulus is still perceived as brightness of a definite intensity. As we are not immediately in a position to compare the sensations of brightness for different colours, this must be done according to the methods of heterochromatic photometry (p. 28). In this way we obtain the results shown in fig. 19, p. 31. As a measure for the *threshold value* of the luminous flux which just gives rise to the sensation of light, we may take the reciprocal of the quantity of light which a point source, e.g. a star, must send to the eye in order that it may just be consciously perceived. This threshold value varies very much according to the circumstances, and in particular depends largely on colour. The least quantity of energy perceptible as light is about $2 \,.\, 10^{-10}$ ergs, provided the eye has been at rest in the dark for a long time previously (dark-adapted eye). It has been found that the yellow spot is by no means the most photosensitive part of the retina. We can convince ourselves of this by trying to pick out a very feeble star in the night sky. It is found that such stars are not seen by *direct* vision, but are seen by *indirect* vision by looking at a point close to but not coinciding with the expected position. That is, the portions of the retina surrounding the yellow spot are more sensitive to very feeble light stimuli than the

* Ger., *Blicklinien.*

yellow spot itself. This peculiarity seems to be due to the absence of rods from the yellow spot of the normal eye; that is, the rods are probably the most photosensitive elements of the eye (Chap. XII, § 2, p. 270).

When the light is feeble, vision takes place wholly by means of the rods, i.e. is indirect; objects appear colourless (*twilight vision*). The greater the stimulated area of the eye, the less the intensity which can still be perceived. For small objects viewed by the rods, the threshold value S is equal to C/φ^2, where C is a constant and φ the angle subtended by the object. On this depends the possibility of improving vision in twilight by means of a telescope, although the latter has the effect of somewhat enfeebling the brightness of extended objects (pp. 129, 136).

3. Accommodation: Spectacles.

Curvature of the Lens.—The most marked refraction to which the incident rays are subject takes place on entering the cornea. The deviation produced by the lens is only trifling. The lens chiefly acts as a correcting organ. To the lens is attached a ring-shaped muscle, the **ciliary muscle,** by which the curvature of the lens may be increased.

This process of increasing the curvature of the lens has been explained by H. Helmholtz and A. Gullstrand. The lens is fixed to two thin membranes. The first is the thin transparent hyaloid membrane, which lies on top of the retina throughout the interior of the eye and encloses the vitreous humour. It merges into the rear part of the lens, to which it is attached along a circle whose plane is normal to the axis of the lens. Between this membrane and the ciliary part of the retina there is inserted a second membrane, the **suspensory ligament.** It is peculiarly folded like a ruff and forms a wavy line along the outer membrane of the lens, being fixed to the lens capsule. It is firmly fixed to the middle (equatorial) parts of the lens; over a great part of the lens the wavy junction-line runs backwards and forwards over the circle of greatest diameter of the lens (equator). The suspensory ligament is also connected to the ciliary processes which run into its folds; these processes are connected to the ciliary muscle. According to Gullstrand, the suspensory ligament is in a state of elastic tension in the eye at rest, and by its elastic pull increases the greatest diameter of the lens (equator), so that the lens, following the pull, is elastically deformed and its front and rear surfaces become flattened. By the contraction of the ciliary muscle the ciliary attachments of the suspensory ligament, especially the parts going to the front surface, are displaced towards the lens. The tension of the suspensory ligament is thereby reduced, its pull on the lens slackens, and hence the elastic deformation diminishes; the lens becomes thicker along its axis, and its surfaces, particularly the front one, become more curved. This relaxation can only go on until the suspensory ligament is slack. The greatest curvature of the lens occurs then (even although the ciliary muscle may contract further).

The mean radius of curvature of the front surface of the lens of a young normal eye (p. 115) varies between 10·4 mm. when the ciliary muscle is not contracted (when the suspensory ligament is fully stretched) and about 5·7 mm. when the suspensory ligament is quite

slack. If the curvature of the lens is considerably increased, the deviation of the rays passing through it is made greater; parallel rays are then reunited in front of the retina, while rays which are divergent on entering the eye can be made to meet on the retina. Hence the eye has the capacity of forming a distinct real image on the retina, even when the object is at a finite distance.

Accommodation.—The power of the eye to alter its focal lengths to suit the distance of the objects observed is called the focusing power or *accommodation* of the eye.

The change in focal length of the lens occurring in the process of accommodation is not, however, entirely due to the change in the curvature and in particular to that of the front surface. On the contrary, the peculiar structure of the lens also plays a part, according to GULLSTRAND'S researches. The substance of the lens, in fact, is not homogeneous, but consists of a large number of very fine membranes lying one above another like the coats of an onion and all contained in the lens capsule. The refractive index of these membranes increases towards the centre. In the process of accommodation, the membranes are subject to relative displacements, the inner layers becoming relatively thicker than the outer ones, and the result is that the mean refractive index of the lens alters to a considerable extent, from 1·409 when the eye is unaccommodated to 1·426 when the accommodation is a maximum.

The Far Point.—A normal eye is capable of making rays from an object at infinity meet on the retina without accommodation, i.e. it gives rise to sharp images of distant objects on the retina. The *far point* of an emmetropic eye, i.e. the most distant point which can still be seen clearly by the eye, therefore lies at infinity. A normal eye is always focused on the far point when accommodation is not being used.

The Near Point.—In the normal eye of a young person the lens may be altered so that it can produce a sharp image on the retina even when the object is as near the eye as 10 cm. The nearest point to which the eye can focus is called the *near point.* Accommodation to very near objects is liable to strain the ciliary muscle, giving rise to pain in the eye. As the eye is accustomed to focus for a distance of 20 to 30 cm. (the distance of the book from the eye in reading or writing), we do not find that the ciliary muscle is strained for this distance. The distance 25 cm. is called the *least distance of distinct vision.*

Far-sightedness.—Eyes whose power of accommodation is so slight that objects can only be seen clearly when they are further than about 30 cm. from the eye, i.e. whose near point is more than 30 cm. from the eye, are said to be *far-sighted.* if the far point is still at infinity. Rays entering a far-sighted eye from an object G less than 30 cm. away (fig. 9) do not meet until they reach a point behind the retina, so that there is a circle of confusion (C) on the retina instead of a point image. The retinal images of objects are then blurred. If a

convergent lens L is put in front of the far-sighted eye, the rays from near objects, which are too divergent, are made less divergent or even parallel, so that they appear to come from a more distant point, for which the far-sighted eye can be focused. For *reading*, the far-sighted person has to wear spectacles with *convex* lenses, their focal length being smaller the more far-sighted he is.

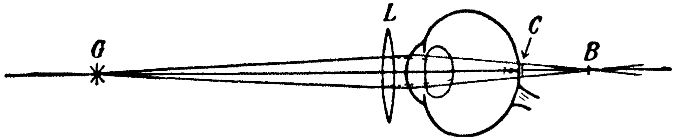


Fig. 9.—Path of the rays in a far-sighted eye with and without spectacles

In old age the eyes become far-sighted and hypermetropic (see below) as a result of the lessened elasticity of the lens. Eyes which have become far-sighted as a result of age are called **presbyopic.*** Hence old people have to wear convex glasses for reading or must hold the page far away from them.

Long-sightedness or **Hypermetropia.**—If an eye can only unite on the retina such rays as are convergent on entering the eye, it is said to be *long-sighted* or *hypermetropic.*† Its far point and its near point are virtual objects lying behind the eye. A hypermetropic eye is altogether incapable of forming a sharp image on the retina of objects in front of the eye. Hence the person with hypermetropia must wear *convex* glasses even for *distant* vision; these will make parallel rays, and also the divergent rays proceeding from objects at a finite distance, converge before entering the eye.

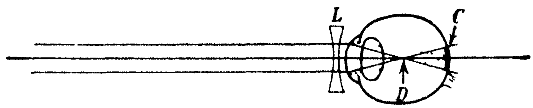


Fig. 10.—Path of the rays in a short-sighted eye with and without spectacles

Short-sightedness or **Myopia.**—If the eye causes parallel rays to meet at a point D in front of the retina (fig. 10), it is said to be *myopic;*‡ there is then a circle of confusion (C) on the retina. Sharp images of objects near the eye are produced on the retina. The *far point* of the myopic eye lies at a finite distance, usually quite close to the eye. In high degrees of myopia the *near point* may lie within a few centimetres of the eye. Short-sighted persons whose far point lies nearer the eye than 25 cm. must use spectacles even for reading. Glasses for *short-sighted* people must be *concave*, in order that the point of meeting

* Gr., *presbys*, old.

† Gr., *hyper*, over.

‡ Gr., *myo*, I close the eyes. Short-sighted people half close the eyes in order to see more clearly (cf. the effect of a stop).

of the rays may be shifted farther back, so that the rays may diverge more strongly on entering the eye, as if they came from a point between the near and far points of the eye.

Short-sightedness cannot be overcome by attempts at accommodation, as the curvature of the lens is always increased, but can never be diminished, by the contraction of the ciliary muscle. If a short-sighted person wears no glasses, serious damage to the eye and even detachment of the retina may result from the attempts of the eye to move the retina nearer the point where the rays meet. The negative focal length of the glasses of short-sighted persons must be smaller the greater the degree of short sight.

Short-sightedness may be due either to too great curvature of the refracting surfaces or to too great length of the eyeball from back to front. It may result from too prolonged use of the eye at short distances (e.g. reading) in youth, owing to the accommodation, usually considerable, which is required.

Dioptric Power.—Spectacle glasses are distinguished by their **refracting power** measured in **dioptres** (p. 65).

The dioptric power of a lens is the reciprocal of its focal length in metres.

For convex lenses it is positive, for concave lenses negative.

Astigmatism.—Many eyes exhibit congenital astigmatism (see Chap. V, § 2, p. 92). This defect may be overcome by means of spectacles with astigmatic lenses. If the eye were normal apart from the astigmatic defect, i.e. if the far point were at infinity for rays in some plane through the axis, the defect could be overcome by a cylindrical lens of a form which would remove the far point in the plane through the axis at right angles to the first to infinity also. As a rule, however, congenital astigmatism is associated with some other defect, e.g. short sight. The spectacles must then consist of lenses differing in curvature in two planes at right angles to one another. Nowadays lenses with toric surfaces are used. A toric surface is the surface obtained by rotating an arc of a circle about an axis which lies in its plane but does not pass through its centre. In order that satisfactory results may follow from the use of astigmatic glasses, it is necessary that the axes of the astigmatism of the eye should first be carefully determined and that then the glasses should be fixed in the spectacle-frame in such a way that their own principal planes of astigmatism coincide with those of the eye.

When ordinary spectacles are worn the astigmatism of oblique pencils (p. 93) has a disturbing effect. The person wearing spectacles does not see objects sharply through them unless he looks through the centres of the lenses in the direction of their axes. If he keeps the spectacles fixed and rotates the eye in its socket, only oblique pencils with astigmatic distortion can pass through the spectacles into the eye. It is very frequently observed that short-sighted persons when looking intently at an object deliberately look obliquely through their glasses, often through the outer margins. This is a sign that their eyes are astigmatic. By experience they have acquired the habit of looking at things obliquely, since in this position the astigmatism of the oblique pencils of light exactly compensates the astigmatism of the eye, or at least reduces it to a minimum. The fact that far more than half of all short-sighted persons have the habit of looking obliquely through their glasses indicates on the one hand that congenital astigmatism is very common. On the other hand, it also indicates that the astigmatism of oblique pencils is very readily perceptible when glasses are used. As a result persons with a high degree of myopia when wearing glasses entirely lose the habit of rotating the eye in its socket; when looking at an extended object they give the eye the required directions by moving the head. This is the explanation of the peculiar fixed look which they have when they take off their

glasses. Hence the recent introduction of glasses which diminish or remove the astigmatism of oblique pencils represents an important advance. The astigmatism may be diminished by meniscus lenses (p. 66), a lens thinner at the centre than at the edges being used for short sight, and a lens thicker at the centre than at the edges for long sight. Such glasses are said to form "punctal" images. The astigmatism of oblique pencils is more completely removed by the so-called punctal glasses,* the surfaces of which do not possess spherical curvature. These non-spherical glasses may be imagined as formed by adding a small portion of the refracting substance to the spherical surface in every direction so as to alter the refraction slightly, and thus remove the astigmatism. When in use, these glasses obviously must always retain the same prescribed position relative to the point about which the eyes rotate.

4. The Visual Angle.

The Visual Acuity or Resolving Power of the Eye.—If in vision only *one* cone is excited, we have the sensation of a *single* luminous point. Hence if the object is so small or so distant that its image does not fall on more than one cone at any one time, we cannot recognize any detail in the object †; the resolving power of the eye is insufficient. The magnitude of the retinal image of an illuminated object depends on the magnitude and distance of the object. The cones are most closely packed together in the yellow spot; here the distance between them is only about 0·004 mm. A retinal image of this magnitude is formed when two rays entering the eye (at the nodal point) include an angle of one minute.‡ This happens, e.g., when rays enter the eye from two points 0·3 mm. apart at a distance of 1 m. from the eye. If a pair of points at a distance of 1 m. are more than 0·3 mm. apart, their image falls on two distinct cones and they are accordingly seen as two separate points (see p. 28). On the other hand, it is still possible to detect the relative displacement of two suitably constructed straight lines when this displacement only amounts to about 10″ (the principle of the vernier).

The Visual Angle.—The angle subtended at the eye by rays from two separate points is called the **visual angle** or **apparent magnitude** (p. 79) of the line joining the two points. The magnitude of the retinal image depends solely on the apparent magnitude. Two objects A and B perpendicular to the optic axis at distances a and b from the eye have the same apparent magnitude if $A/a = B/b$. This ratio is the tangent of the visual angle.

In vision with *one* eye (for the phenomena of binocular vision, see Chap. XII, § 4, p. 276) we can only form a judgment about the value of the visual angle, i.e. of the ratio of the actual magnitude of an object and its distance; hence we cannot deduce its actual magnitude unless we know its distance, or its actual distance unless we know its magnitude. If we know neither of these quantities

* First calculated by M. VON ROHR for the Carl Zeiss optical works in Jena.

† EUCLID (300 B.C.) noticed that it appears "round".

‡ The discovery that the resolving power reaches its limit for a visual angle of 1′ was made by HOOKE in 1674.

by experience, we may be completely deceived as to the actual magnitude and actual distance of the object.

As the tangent of the visual angle of an object is the quotient of its magnitude and its distance, the visual angle may be increased in two ways, either by increasing the size of the object or by diminishing its distance from the eye. Bringing the object nearer the eye has the same effect on the size of the retinal image as increasing the size of the object. If, then, we wish to see the details of an object, we must bring it so close to the eye that the visual angle associated with the details to be distinguished exceeds one minute. This process, however, is subject to the limitations imposed by the refracting parts of the eye: an object may not be brought nearer than the near point of the eye, otherwise its image will be blurred.

5. The Hand-lens or Magnifying-glass.

Object Nearer the Eye than the Near Point.—Let the point G (fig. 11) lie on the optic axis between the eye and its near point. A ray of light starting from G (shown dotted) meets the surface representing the spherical surface of the reduced eye (§ 2, p. 115) at C.

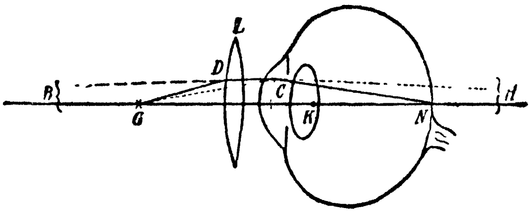


Fig. 11.—Path of the rays when a hand lens is used

Owing to the too great divergence of the light rays coming from G the refracted ray only cuts the optic axis at a point H lying far behind the retina. CH may even diverge from the optic axis. By means of a convex lens L placed immediately in front of the eye, the path of the ray may be so altered that it falls on the yellow spot N.

Even a ray GD leaving the point G at a greater angle to the axis will reach the yellow spot by the path GDCN. The ray DC entering the eye is inclined at an angle which appears to be that of a ray starting from the point B. If B is the near point of the eye, the eye in conjunction with the lens L can still see the point G as a point (without a circle of confusion being formed). The eye without the lens L would not see the object sharply unless it were moved back to B.

Magnified Image at the Near Point.—Let an extended object have the actual magnitude denoted by GG′ or BG″ (fig. 12). If the object is at G, it subtends the angle $\psi = \mathrm{GKG}'$ at the eye; if the object is at B, the angle is $\phi = \mathrm{BKG}''$. The size of the retinal image is proportional to the tangent of the visual angle, which in future we shall replace by the angle itself measured in radians, as the angles used are small.*

* The deductions made here are true only for rays inclined at a small angle to the axes.

The subjective magnification w (p. 79) is determined by the ratio of the two visual angles ψ and φ; hence

$$w = \frac{\psi}{\varphi} = \frac{GG'/KG}{BG''/KB} = \frac{KB}{KG} = \frac{v}{u}.$$

Here KG is the true distance between the eye and the object (u), while KB is the distance between the eye and the virtual image produced by the convex lens L (v). If we now assume that the lens L is held so close to the eye that we

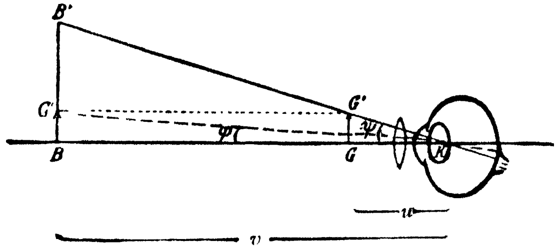


Fig. 12.—Magnification by a hand lens

may regard the nodal point of the lens as coinciding with that of the eye, the two quantities u and v, both of which are to be taken as negative here, are related to f_1, the focal length of the lens, by the general lens equation (p. 41) $-1/v = -1/u + 1/f_1$, whence $v/u = 1 + v/f_1$.

The subjective magnification produced by the hand-lens is therefore

$$w = 1 + \frac{v}{f_1}.$$

If we hold the hand-lens immediately in front of the eye and look at the object through it, using the greatest accommodation possible, we bring the object to a finite distance such that the image BB′ lies at the near point of the eye. If we call the distance of the near point from the eye d, the magnification is

$$w = 1 + \frac{d}{f_1}.$$

Hence it follows that when a hand-lens is used in conjunction with the fully-accommodated eye, its magnifying effect is greater the smaller its focal length and the greater the distance of the near point of the eye from the eye. The hand-lens is therefore most useful to far-sighted persons; in the case of short-sighted persons, for whom d is small, the magnification is only trifling.

Magnified Image at the Far Point.—Frequently it is necessary to consider the process of magnification by a hand-lens from a somewhat different point of view. Let an object BG″ (fig. 12) be at the distance BK $= v$ from the nodal point of the reduced eye. Its apparent magnitude ϕ is given by BG″$/v$. If its image on the retina is of magnitude

y', we have $\phi = y'/l$, where l is the distance of the point K from the retina. Now suppose a magnifying-glass is interposed between the object BG″ and the eye. Let the object be moved away until the eye can see its image in the magnifying-glass without any accommodation.

As a matter of fact all optical instruments are most effectively used with the *unaccommodated* eye. This is true of the microscope and telescope as well as of the magnifying-glass. According to A. GULLSTRAND, however, inexperienced users, and especially children, are apt to accommodate for near vision quite unnecessarily when they use these optical instruments.

The magnifying-glass must then make every point of the object give rise to a parallel pencil of rays, so that the image produced by the magnifying-glass may lie at infinity. If this is to be the case, however, the object GG′ must lie in the focal plane of the magnifying-glass away from the eye, i.e. at a distance f from the eye. If from G′, the point of the object farthest from the axis, we draw the ray through the first principal point (through the centre of an infinitely thin magnifying-glass) its inclination to the axis is ψ, where $\psi = G'G/f$. Hence in the image space of the lens this is the inclination of the whole pencil of rays which diverges from G and reaches the eye as a parallel pencil. The object accordingly has the apparent magnitude ψ. An image of magnitude y'' is thrown on the retina of the *unaccommodated* eye. Then $\psi = y''/l$, if we assume that the change of accommodation from that for the distance v to that for the far point has not noticeably altered the distance of K from the back of the eye. In looking twice at the object, first at the distance v from the eye and without the glass, and then with the glass, the eye being at its focus G, we are conscious of the magnification given by

$$w = \frac{y''}{y'} = \frac{l \tan\psi}{l \tan\phi} = \frac{GG'}{f} \bigg/ \frac{BG''}{v} = \frac{v}{f}.$$

If we replace $1/f$ by D, the refracting power of the magnifying-glass in dioptres, we obtain

$$w = v\mathbf{D}.$$

The magnification produced by a hand-lens when used with the unaccommodated eye is the product of the distance of the object when looked at without the lens and the refracting power of the lens.

Here v is to be measured in metres if D is measured in dioptres.

The factor v is a perfectly arbitrary quantity and depends on the distance of the object from the eye before the lens was used, or in other words, what image was used as a comparison in reckoning the magnification (*individual magnification*). It is conventional to use the "least distance of clear vision", taken as 25 cm., for v. This being settled once and for all, we should have $w =$ D/4 (*conventional magnification*); the magnification would then be always equal to a quarter of the refracting power of the lens.

"Verant" Lenses.—Hitherto we have assumed that the magnifying-glass is being used with the eye at rest. This applies to strong magnifying-glasses. These can only be corrected for regions near the axis and give sharp images of such regions only. Hence we involuntarily look through the lens in the direction of the axis; if we wish to see another point of the object clearly, we direct both eye and lens at that point without altering their relative positions. In this case the central rays of the pencils from various points of the object meet at the centre of the pupil of the eye.

If, however, the magnifying-glass is to give a wider field of view, which of course can only be attained for small magnifications, this assumption no longer suffices. In looking over the image of the object, the eye then rotates in its socket. The central rays must therefore intersect at the point about which the eye rotates. Hence the spherical aberration and chromatic aberration (p. 167) of a lens used in this way may be got rid of by ensuring that in every position the eye receives pencils of rays free from distortion and astigmatism and in the correct direction. The necessity of such correction relative to the point about which the eye rotates was first pointed out by A. GULLSTRAND. Glasses fulfilling this condition were calculated by M. VON ROHR in the Zeiss works at Jena and made available commercially under the name of *Verant lenses*.

Centre of Perspective of a Photograph.—As the point about which the eye rotates is the centre of perspective for our three-dimensional vision (p. 116), a picture does not give us the proper impression of space unless the central rays from the picture meet at the point about which the eye rotates and are inclined at the same angles to one another as if we were looking directly at the object depicted. Otherwise the picture appears distorted or the relative magnitudes seem to be wrong. The image which a photographic lens forms of a very distant object lies in its focal plane. The centre of perspective is the first focus; for if the centre about which the eye rotates lies at the first focus of the photographic lens, the eye sees every image-point of the second focal plane through the lens exactly at the same distance and by rays making the same angle with the axis behind the lens as the object originally photographed did in front of the lens. Hence if we are to gain a natural impression from a photograph, the picture should be looked at through a lens with the same focal length as the camera lens. If we look at the picture in any other way with the naked eye the picture will not seem correct unless we can imitate the effect of the lens by suitably accommodating the eye; otherwise the impression must always be impaired owing to the relative magnitudes being incorrect. It is in fact very surprising how well a photograph stands out even when looked at by a single eye when this is assisted by an ordinary magnifying-glass. For examining photographs, Verant lenses of suitable focal length are specially made; when these are used the picture must be at one focus and the point about which the eye rotates at the other. The same result is of course obtained if a suitably enlarged proof is made of a photograph and examined by the naked eye.

6. Telescopes.

The object of telescopes or telescopic systems (see p. 88) is to increase the size of the retinal image of a distant object, i.e. to increase the angle subtended by the object at the eye, without altering the accommodation of the eye. As an object at an infinite distance sends parallel rays to the eye a normal eye sees the object without accommodation, whether a telescope is used or not (accommodation for infinity). It follows that rays which are parallel on entering the telescope must still be parallel when they leave it.

The two main types of telescopic system are (1) the **Galilean telescope,** which consists of a convergent objective (with a positive focal length) and a divergent eyepiece (with a negative focal length), and (2) the **astronomical telescope,** which consists of a convergent objective and a convergent eyepiece.

The Galilean Telescope. (*a*) *Construction.*—In the Galilean * telescope (fig. 13, Plate XI †) a convergent objective L_1 and a divergent eyepiece L_2 are combined in such a way that the second focus of the objective coincides with the second focus of the eyepiece.

The convergent objective L_1 has the focus F. The light rays entering the objective parallel to its principal axis are refracted so as to converge at the focus F. The divergent eyepiece L_2, which is to render the rays parallel again, has its second focus at F. It follows that l, the distance of the eyepiece from the objective, i.e. the length of the Galilean telescope, must be equal to the difference of the focal lengths of the objective and the eyepiece:

$$l = f_1 - f_2.$$

(*b*) *Path of the Rays.*—The path of the rays in the Galilean telescope for any ray (not parallel to the axis) is shown in fig. 14. The two lenses, the convergent objective and the divergent eyepiece, are placed coaxially. Their common focal plane is F_1F_1.

If an observer with the eye at rest looks through a Galilean telescope, only those rays are concerned in the formation of the image which enter P, the pupil of the eye, the latter being immediately behind the eyepiece.

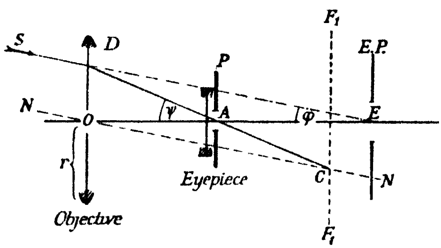


Fig. 14.—To illustrate the path of the rays in the Galilean telescope

* Galileo was not the first to construct a telescope of the type now known by his name. The discovery is ascribed to one or other of the Dutch spectacle-makers, Zacharias Jansen (1604), or Franz Lippershey of Middelburg. The latter applied for a Dutch monopoly on 2nd October, 1608. Galileo heard of the discovery and thereupon made experiments on the combination of two lenses; in May, 1609, he succeeded in constructing a telescope. Galileo's chief contribution was the *improvement* of the telescope, as those made by him were far superior to the others available at the time. They enabled detailed astronomical observations to be made; these attracted an immense amount of attention and led to extremely important conclusions about the heavenly bodies.

† Fig. 13 was obtained by photographing the path of the rays as follows: first the paths of the rays were photographed as produced by the objective L_1 alone, i.e. without the eyepiece L_2; another exposure was then made on the same plate after the eyepiece L_2 had been inserted, making the rays which converged to F in the former exposure parallel again.

Only those rays which after refraction pass through A, the centre of the pupil, can be central rays. Let DA be such a ray. As it also passes through the centre of the eyepiece, it suffers no change of direction. If prolonged further, the ray DA would meet the focal plane F_1F_1 at C. If we draw the subsidiary axis NN passing through C and the centre of the objective, we know that the ray DA before entering the objective must travel parallel to this axis (as C, its point of intersection with the subsidiary axis NN after refraction, lies in the focal plane F_1F_1). We thus obtain the direction of the incident ray S by drawing SD parallel to NN. If the objective were not there, SD would go on as a straight line and meet the principal axis at E; hence E must lie in the entrance pupil EP. Thus we may summarize the foregoing as follows:

The prolongation of the ray S meeting the objective intersects the axis of the telescope in the entrance pupil E at the angle ϕ. This is the angle which the distant object from which the ray S comes would subtend at the eye if there were no telescope present. The ray DA which travels through the telescope and enters the eye intersects the axis of the telescope in the exit pupil A at the angle ψ. This is the angle subtended at the eye by the same object when the latter is observed through the telescope.*

(c) *Magnification.*—The magnification of the visual angle is (p. 79) the subjective magnification $w = \frac{\psi}{\phi} = \frac{DO/OA}{DO/OE} = \frac{OE}{OA}$. Here OA is the length of the telescope, $l = f_1 - f_2$. As E is the centre of the entrance pupil, A is its virtual image produced by the objective; hence we have the equation $1/OA = 1/OE + 1/f_1$. It follows that $\frac{OE}{OA} = \frac{f_1}{f_1 - OA}$, and as $OA = f_1 - f_2$, we have

$$w = \frac{\psi}{\phi} = \frac{f_1}{f_2}.$$

In the Galilean telescope the magnification is equal to the quotient of the focal lengths of the objective and the eyepiece.

(d) *The Field of View.*—The field of view on the image side is limited (in the case of an eye at rest) by the rays which enter by the outermost edge of the objective and pass through A, the centre of the pupil of the eye. *The objective is the field stop of the Galilean telescope.* Fig. 15 (Plate XII) is from a photograph of the paths of an axial pencil of parallel rays and of a pencil of parallel rays passing through the edge of the objective.

If the radius of the mount of the objective is r, the angle determining the field of view on the image side is given by $\psi = r/(f_1 - f_2)$,

* These statements are subject to alteration if the eye is moved; the central rays then intersect not at the centre of the pupil but at the point about which the eye rotates.

by fig. 14. The angle determining the field of view on the object side is accordingly

$$\phi = \frac{\psi}{w} = \frac{1}{w} \cdot \frac{r}{f_1 - f_2} = \frac{f_2}{f_1} \frac{r}{f_1 - f_2}.$$

From this equation it follows that *the field of view on the object side diminishes in the same ratio as the magnification increases, the length of the telescope remaining the same.*

In actuality the calculated field of view cannot be completely utilized. If while using the telescope we move the eye about, the rays which limit the field of view intersect not at the centre of the pupil but at the point farther back about which the eye rotates. This reduces the maximum aperture of the field of view on the image side (2ψ) in fig. 14 and hence also the maximum aperture of the field of view on the object side (2ϕ).

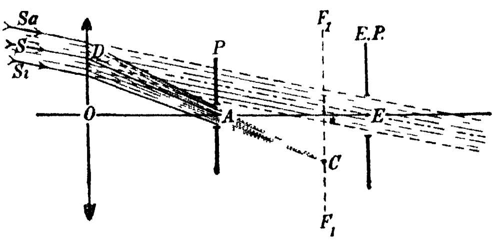


Fig. 16.—The entrance pupil EP and the exit pupil AP of the Galilean telescope

(*e*) *The Brightness of the Image.*—In the formation of the image of any point, only that pencil of rays is concerned which after refraction enters the pupil P (fig. 16). If the telescope were not there this pencil would pass through the entrance pupil as a pencil of parallel rays. Hence the brightness of the image of a single point is determined by the size of the entrance pupil. The radii of the entrance pupil and of the pupil of the eye (ρ, p respectively) are connected by the relationship $\rho/p = CD/CA = f_1/f_2$. As $f_1/f_2 = w$, it follows that

$$\rho = pw.$$

B, the brightness of the telescopic image of a single point, is proportional to the luminous flux reaching the eye from the object and hence to the area of the entrance pupil EP, i.e. proportional to $\pi\rho^2$; hence the brightness of the image may be written

$$B \propto \pi p^2 w^2.$$

The pupil of the eye acts as an aperture stop when the Galilean telescope is used. The brightness of the image of a point is proportional to the square of the radius of the pupil of the eye and the square of the magnification produced by the telescope.

The Galilean telescope increases the *visual angle* and the *brightness of single points of the image* simultaneously. If we look at the night sky through an opera-glass (which consists of two Galilean telescopes side by side), we see a considerably

larger number of fixed stars than we do with the naked eye. This is not due to the visual angle of a star being increased, as it is still infinitely small, but is due to the increase in brightness. As the magnification of the visual angle has the effect of moving the individual points farther apart, however, there is no increase in the *surface brightness* of the field of view or of an extended object. The trifling losses occurring in the passage of the light through the lenses of the telescope are of no great consequence. Hence the brightness of the image produced by the Galilean telescope is one of its great advantages. Accordingly it is frequently selected for use in observing feebly illuminated objects, e.g. as a night-glass or an opera-glass (see also p. 117).

The Astronomical Telescope.—In the astronomical telescope * (fig. 17) the two convergent lenses, the objective and the eyepiece, are set coaxially in such a way that the second focus of the objective coincides with the first focus of the eyepiece. As a result the rays SS parallel to the axis converge to F after refraction at the objective and, diverging from F, impinge on the eyepiece. Here they are again refracted, in such a way as to leave the eyepiece parallel to the axis.

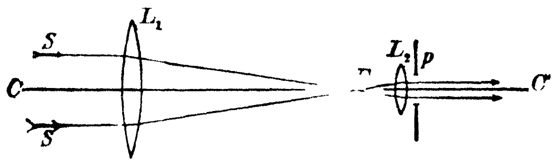


Fig. 17.—Path of the rays in the astronomical telescope (diagrammatic)

The purpose of the objective is to form a real image of a distant object in its focal plane. This image is then observed through the eyepiece, the latter acting as a magnifying-glass. Thus if the image on the focusing screen of a camera is looked at through a magnifying-glass, the combination of instruments represents an astronomical telescope.

In the astronomical telescope the distance of the eyepiece from the objective, i.e. the length of the telescope (l), is equal to the sum of the focal lengths of the objective and the eyepiece:

$$l = f_1 + f_2.$$

(*a*) *Aperture Stop.*—In actual astronomical telescopes the objective invariably has a long focal length and the eyepiece a short focal length, so that in our discussion we shall assume this to be the case. From fig. 17 we see at once that the diameter of the incident pencil of parallel rays SS bears the same ratio to the diameter of the pencil of rays emerging from the eyepiece of the astronomical telescope as the focal

* The astronomical telescope, which consists of two convex lenses, is described by KEPLER in a treatise on optics which appeared in 1611, but KEPLER himself never made or used it. It was first constructed by the Jesuit father CHRISTIAN SCHEINER in 1615. SCHEINER also pointed out that by using three lenses a telescope can be constructed which is suited to the observation of terrestrial objects, the image being then right side up ("terrestrial" telescope, p. 133).

length of the objective does to the focal length of the eyepiece. If the apertures of the lenses are to be fully utilized, therefore, the diameters of the objective and the eyepiece of a properly constructed telescope must be in the ratio of the focal lengths, i.e. the two lenses must have the same relative aperture (p. 110).

The aperture of the eyepiece must not greatly exceed the diameter of the pupil of the eye, otherwise some of the rays will not enter the eye. If, for example, we assume that the focal length of the objective is ten times that of the eyepiece, the diameter of the emergent pencil of rays is only a tenth of that of the incident pencil of rays. Let p, the diameter of the pupil of the eye, be 6 mm. Then if the pupil and the mount of the objective are to limit the incident pencil to the same extent, the diameter of the objective must be 6 cm. In actuality the telescope is usually so constructed that the diameter of the objective mount is smaller than the ratio stated above. It follows that in this case the quantity of light entering the telescope is determined by the mount of the objective alone.

The mount of the objective usually forms the aperture stop of the astronomical telescope.

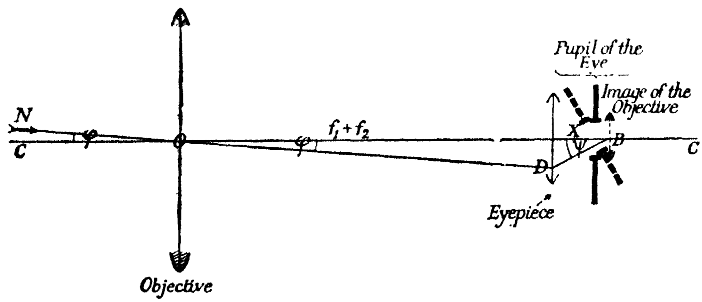


Fig. 18.—Magnification produced by the astronomical telescope

(*b*) *Exit Pupil.*—In the case which usually occurs in practice (fig. 18), the central rays from all points of the object, even those farthest to either side, pass through the centre of the objective; they are then refracted towards the axis by the eyepiece and intersect the axis where the image of the objective formed by the eyepiece is situated.

The optical image of the objective is therefore the actual exit pupil of the telescope.

In some astronomical telescopes which give an image of great intensity (night-glasses), this exit pupil may be larger than the pupil of the eye. In this case the pupil of the eye is the aperture stop. If the eye were *at rest*, i.e. if indirect vision were used, a telescope would be completely utilized if the image of the objective aperture produced by the eyepiece fell on the pupil of the eye. But as we examine the field of view the eye rotates in its socket. Hence if the eye is moved all the central rays do not enter the eye unless the exit pupil of the telescope lies at the point about which the eye rotates; the exit pupil must then be greater than the pupil of the eye.

(*c*) *Magnification.* The relationships are illustrated in fig. 18. NO is the central ray from a point on one side of the object through the centre of the objective (it is at the same time a subsidiary axis of the objective). The prolongation of NO meets the eyepiece at D and is refracted to the point B on the principal axis. Let B (the point about which the eye rotates) be at a distance x from the eyepiece. The incident rays form the visual angle NOC $= \varphi$ with the axis of the telescope; when looked at through the telescope the same point of the object subtends the angle DBO $= \psi$. The ratio of the magnitude of the retinal images with and without the telescope is the subjective magnification w (p. 79) of the telescope.

From fig. 18 it follows immediately that

$$w = \frac{\psi}{\phi} = -\frac{f_1 + f_2}{x}.$$

As B is the real optical image of the objective, produced by the eyepiece of focal length f_2, we have the general lens equation (p. 67)

$$\frac{1}{x} = -\frac{1}{f_1 + f_2} + \frac{1}{f_2},$$

whence
$$\frac{f_1 + f_2}{x} = \frac{f_1}{f_2};$$

hence
$$w = -\frac{f_1}{f_2}.$$

The magnification of the astronomical telescope is equal to the quotient of the focal lengths of the objective and the eyepiece.

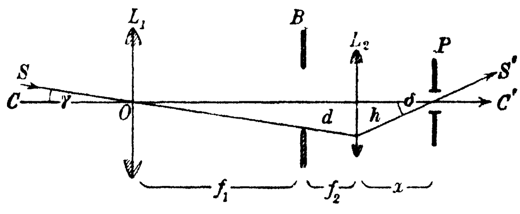


Fig. 19.—Field of view on the object side in the astronomical telescope

The Field Stop.—In the common focal plane we set up a material screen, which lets through only the central rays. As a real image of the infinitely distant object is produced in the common focal plane, the screen set up here limits the field of view. *It is the actual field stop of the astronomical telescope.*

The field of view on the object side is limited by that pencil of rays whose central ray S (fig. 19) passes through O, the centre of the objective L_1, and grazes

the edge of the stop B. If the radius of the opening of the stop is d, the visual angle on the object side is determined by

$$\gamma = \frac{d}{f_1}.$$

The area of the field of view is proportional to the square of this expression. Hence *the field of view on the object side in the astronomical telescope is directly proportional to the aperture of the stop and inversely proportional to the square of the focal length of the objective; it is independent of the focal length of the eyepiece and of the aperture of the objective.*

The visual angle (δ) on the image side (fig. 19) is determined by the magnification:

$$\frac{\delta}{\gamma} = w = \frac{f_1}{f_2}.$$

Hence

$$\delta = -\frac{f_1}{f_2}\gamma = -\frac{d}{f_2}.$$

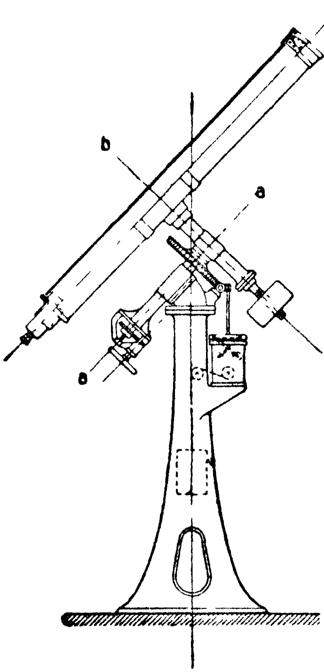


Fig. 20.—Telescope with equatorial mounting

The field of view on the image side in the astronomical telescope is directly proportional to the aperture of the stop and inversely proportional to the square of the focal length of the eyepiece.

(In the Galilean telescope it is the *diameter of the objective* that determines the field of view.)

"Peephole" Observation.—If the exit pupil of the astronomical telescope can be got to coincide with the point about which the eye rotates, the whole field of view can be utilized even when the eye is moved, whereas in the Galilean telescope moving the eye restricts the aperture of the field of view. The construction of the instrument, however, does not always permit one to place the eye near the real image of the objective mount produced by the eyepiece; the exit pupil then lies in front of the point about which the eye rotates. The field of view of the telescope can in this case only be looked at through this little circle, as if through a window; the field of view is limited, as it were, by a peephole. To examine as large a field of view as possible, it is useless to rotate the eye in the socket; on the contrary, one must move the head and the eye from side to side in order to catch rays which pass through the hole obliquely.

Position of the Image.—Fig. 18 shows that the angles ϕ and ψ are on *opposite* sides of the axis; hence it follows that an object looked at through the astronomical telescope appears *inverted*.

Equatorial Mounting of Telescopes.—In observatories telescopes are usually set up as shown in fig. 20. The telescope can be rotated about the axes *aa* and

bb, and hence can be pointed in any direction (this being determined by declination and right ascension (Vol. 1, p. 26)). The axis *aa* is parallel to the earth's axis and is called the polar axis of the telescope. To this axis is attached clockwork which rotates it once in every twenty-four (sidereal) hours in the direction opposite to that in which the earth rotates, i.e. rotates the axis in the direction in which the fixed stars rotate. Then if the telescope is focused on a particular star, that star continues to be visible in the telescope (*equatorial mounting*).

The Terrestrial Telescope.* To get rid of the inversion of the image in the astronomical telescope, which is often a troublesome drawback, another convergent lens (the *erecting lens*) is inserted between the objective and the eyepiece (fig. 21).

The objective causes the parallel rays *SSS* entering the telescope to meet at the point A of the focal plane passing through F_1, the focus of the objective. From here they diverge and are again made to converge at B by the erecting

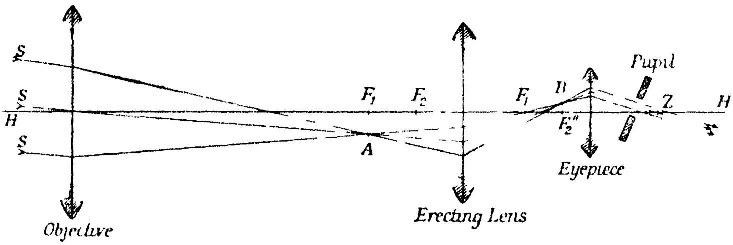


Fig. 21.—Path of the rays in the terrestrial telescope (diagrammatic)

lens. If the distance of the focus F_1 of the objective from the erecting lens is double the focal length of the erecting lens (as is assumed in the figure) the point B also lies at a distance from the erecting lens equal to twice its focal length, on the other side of the axis. From here the rays again diverge and fall on the eyepiece. If B lies in the first focal plane of the eyepiece, the rays are parallel when they leave the eyepiece and enter the pupil of the eye. (Z is the point about which the eye rotates.)

The rays then enter the eye from the *same* side of the axis of the telescope as they would if the object were observed without the telescope; when looked at through the telescope the object now appears *right way up*.

In the case specially discussed here, where the point B is at the same distance from the erecting lens as the point A, the distance of the focal plane of the objective from the focal plane of the eyepiece is equal to four times the focal length of the erecting lens. The astronomical telescope must therefore be lengthened by this amount if it is to be transformed into a terrestrial telescope by the addition of an erecting lens.

If f is the focal length of the objective, f' the focal length of the

* Described by Kepler about 1611.

erecting lens, and f'' the focal length of the eyepiece, l, the length of the terrestrial telescope, is

$$l = f + 4f' + f''.$$

Hence a terrestrial telescope is always inconveniently long.

The terrestrial telescope is also subject to a further disadvantage. The rays passing through the centre of the objective do not pass through the centre of the erecting lens; hence the erecting lens may itself act as a field stop and diminish the field of view. Finally, the rays refracted through its margin are responsible for an increase in the spherical aberration and distortion.

Note. The above illustration of the structure of a terrestrial telescope is diagrammatic only; actual terrestrial telescopes differ from it in many ways. If we arrange two astronomical telescopes end to end so that the exit pupil of the first coincides with the entrance pupil of the second, this combination also gives an erecting telescope, the reversal of the image being brought about by the eyepiece of the first telescope and the objective of the second.

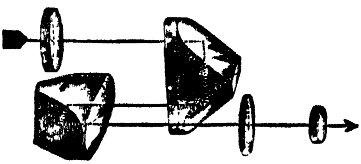


Fig. 22.—Path of the rays in the prismatic telescope

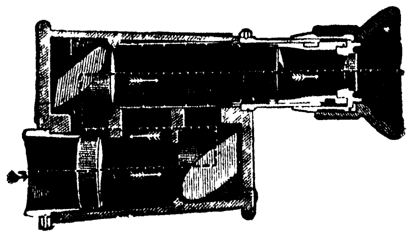


Fig. 23.—Abbe's prismatic telescope

The Prismatic Telescope.*—The image may also be made erect by means of totally reflecting prisms (fig. 22), which at the same time enable the telescope to be made considerably shorter (fig. 23).

The rays entering the objective in the direction shown by the arrow are totally reflected at the two side faces of a right-angled prism; their direction is reversed, and they return parallel to their original direction and fall on a second totally reflecting prism, whose refracting edge is at right angles to that of the first prism. Here the direction of the rays is again reversed by double reflection; they now proceed in their original direction, but displaced to one side, and pass through the eyepiece of the telescope.

The two reflections in the first prism alter the relative positions of the rays in such a way that rays which are originally above the principal axis and parallel to it subsequently lie below it. The reflections in the second prism similarly cause an interchange of left and right. The double exchange exactly reverses the relative positions of the rays; hence the image of the object, which without

* The use of erecting prisms was discovered as early as 1850 by Ignazio Porro—an Italian engineer who resided chiefly in Paris, and soon afterwards attempts were made in Paris to make prismatic telescopes, but were given up owing to failures due to faulty melting of the glass and insufficient accuracy of polishing. Thus the discovery was soon completely forgotten. In 1893 E. Abbe reinvented the prismatic telescope, and the first prismatic telescopes of practical use were made in the Zeiss works. Abbe first heard of Porro's attempts through the German patent office, which brought them back to the light of day and on their account refused Abbe a patent.

the two reflecting prisms would appear inverted, is reproduced in such a way as to correspond exactly with the object as regards both up-and-down and left-and-right.

In view of its shortness, the prismatic telescope is often made in duplicate for use by both eyes at once (binoculars), and is employed for opera-glasses or field-glasses.

As the rays entering the telescope are laterally displaced, the objectives may be placed farther apart than the eyepieces. The result is that an observer can look at a landscape through prism binoculars as if his eyes were farther apart. This makes the relief of the landscape stand out more clearly (Chap. XII, § 4, p. 278). In the telestereoscope or relief telescope the distance between the objectives is made particularly large. In the *periscopes* used in submarines or in trench warfare the folding-up of the path of the rays by means of totally reflecting prisms is used to bring the objective to a convenient position for observation.

Note.—In other applications of the same principle the two prisms are replaced by a single "erecting prism" of special form; the prismatic telescopes of many makers contain silver mirrors instead of totally reflecting prisms.

The Reflecting Telescope.—The objective of the astronomical telescope may also be replaced by a concave mirror. As an example of one of the numerous possible types of construction, fig. 24 shows a type due to NEWTON. The rays reflected by the concave mirror are displaced to one side by a small plane mirror, and the image produced by the latter is observed through the eyepiece. The plane mirror is so small that it does not materially impair the brightness of the image; peculiar diffraction effects, however, are sometimes produced by its mount (fig. 36, Plate XII).

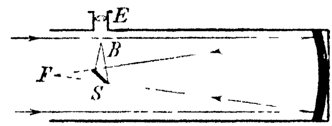


Fig. 24.—Principle of Newton's reflecting telescope

Although the importance of the reflecting telescope was somewhat diminished by the invention of achromatic lenses, at the present time it is one of the astronomer's most powerful aids to investigation, now that it is possible to produce and mount large-size mirrors of the extraordinary degree of precision which is necessary. On Mount Wilson in California there is a reflecting telescope of diameter 258 cm. and focal length 12·9 m., by means of which important new results bearing on the structure of the universe have already been obtained (Vol. V). A reflecting telescope about 7 m. in diameter is under construction at the same place.

Applications of the Telescope. (*a*) *Use of the Telescope in Astronomy.*—In order to understand how the telescope acts when a distant object is observed through it, we must distinguish between the two cases: (1) when the object, though very distant, still subtends a measurable angle at the eye, (2) when as a result of extreme remoteness of the object the angle subtended at the eye falls below the value given in § 4, p. 121, even with the strongest magnification attainable, i.e. the object appears as a single point. The latter case occurs e.g. when we observe a fixed star through a telescope (p. 129). Both to the naked

eye and through a telescope a fixed star appears as a point, or rather as a minute disc surrounded by diffraction rings (p. 145), so that there can be no question of actual magnification of a single fixed star. Nevertheless, a telescope does serve a useful purpose in the observation of fixed stars; it magnifies the distances between stars and increases their apparent brightness.

When a fixed star is observed with the naked eye, only that cone of rays enters the eye which has the opening of the pupil as base (the pupil is the aperture stop); whereas when an accurately constructed astronomical telescope is used, the retinal image is produced by the whole cone of rays which has the objective as base (the objective is the aperture stop). Hence point objects are brighter in the ratio which the size of the objective bears to the size of the pupil of the eye; in a correctly constructed telescope this ratio is the square of the magnification. The uniformly bright background, or an extended object, can at most reach its natural brightness (p. 129). Hence with the telescope we can see fixed stars which owing to their feeble intensity are no longer perceptible to the eye, but still as points of light only. The same is true for the reflecting telescope.

It is for this reason that telescopes with objectives of very great diameter are used in observatories (see p. 147). In order that the spherical aberration due to rays entering the marginal regions of the objective may not be excessive, the objective must not be strongly curved, i.e. it must have a very long focal length. The two reasons are together responsible for the giant dimensions of the telescopes used in modern observatories. The objectives of the four largest refractors in the world have the following dimensions: Yerkes Observatory, Williams Bay, Wisconsin, diameter 102 cm., focal length 19 m., Lick Observatory, Mount Hamilton, California, diameter 91 cm., focal length 18 m., Meudon Observatory, near Paris, diameter 83 cm., focal length 16 m., Astrophysical Observatory, Paris, diameter 80 cm., focal length 12 m. See also above.

In addition to the brightness, the **resolving power** of the telescope (p. 121) increases as the diameter of the objective is increased. Hence the larger the diameter of the objective, the greater the ease with which the two constituents of a double star can be observed separately. It is found in practice that the least angular distance between two stars which are still separately distinguishable is $\varphi = 116''/d$, where d is the diameter of the objective in millimetres. The critical value $d/116''$ is taken as a measure of the resolving power. For an interference method of obtaining very high resolving power see pp. 201–202.

The use of very high magnification is subject to a good many limitations. The reason lies in the unsteadiness of the air, the irregular mingling of warmer and cold currents, which causes the ray of light to be bent irregularly here and there, so that the objects under observation appear to move about. The irregular sparkling or scintillation of the fixed stars in the night sky, which is visible even to the naked eye, arises in the same way and serves to illustrate the disturbing effects (cf. p. 261). The largest modern telescopes are therefore erected on mountains in woodland areas. The magnification of a telescope is also limited to a certain value (about 1000 diameters) by unavoidable errors of aberration.

(*b*) *Use of the Telescope to observe Objects near at hand.*—Hitherto we have assumed that the objects looked at through the telescope

are so remote that the rays entering the telescope from a point of the object may be regarded as parallel. Now telescopes may also be used to observe objects at a finite distance. In this case the distance between the eyepiece and the objective must be increased so that the real image of the object produced by the objective coincides with the first focus of the eyepiece when the eye is focused for infinity, i.e. is used without accommodation. Hence the telescope must be pulled out farther the nearer the object is.

(*c*) *Cross-wires, Graticules.*—In telescopes used for quantitative work **cross-wires** are placed at the point where the real image of the object is situated; to an observer looking through the telescope, these appear to coincide with the object. The cross-wires consist either of two fine threads (from the egg-cocoon of the garden spider), one vertical and one horizontal, intersecting exactly on the axis of the telescope, or of a system consisting of several horizontal and vertical threads. The cross-wires may be replaced by a scale produced by engraving on glass with a diamond-point, etching, or photography, by means of which the magnitude of the visual angle may immediately be read off against the separate parts of the object. Such scales are known as **graticules**, or, in this particular case, as eyepiece micrometers (see also Vol. I, p. 15).

7. The Microscope.

Construction.—The microscope, like the astronomical telescope, consists of a convergent objective and a convergent eyepiece. The essential difference between them is that the rays entering the objective do not start from an object at an infinite distance, i.e. do not enter as parallel rays. The object lies near the first focus of the objective at a distance greater than the focal length. The second focus of the objective and the first focus of the eyepiece do *not* coincide as they do in the Galilean telescope, but are at a definite distance Δ apart (the optical tube-length (p. 86)).

The purpose of the objective is to produce a real and highly magnified image of an object placed very near its focus. This image is viewed through the eyepiece, which acts as a magnifying-glass. If, therefore, the image thrown on a screen by a lantern is looked at from the back through a magnifying-glass, the combination of projection lantern and magnifying-glass is essentially a microscope.

In fig. 25 O is the centre of the objective and F_2F_2, F_1F_1 are its focal planes; O′ is the centre of the eyepiece and $F_2'F_2'$, $F_1'F_1'$ are its focal planes. Let the focal length of the objective be f_1 and the length of the eyepiece be f_2. The object G is at a distance from the objective greater than its focal length.

Path of the Rays.—We shall trace out the path of an arbitrary ray starting from an object-point G which does not lie on the principal axis CC′. To determine the direction of the refracted ray, we draw

the subsidiary axis NN parallel to GD. The ray meeting the objective at D passes through the intersection of this subsidiary axis with the focal plane F_1F_1 and meets the eyepiece at E. If we then draw the subsidiary axis N′N′ through O′, the centre of the eyepiece, parallel to DE, and join its point of intersection with the focal plane $F_1'F_1'$ to E, we obtain the refracted ray, which cuts the principal axis at J.

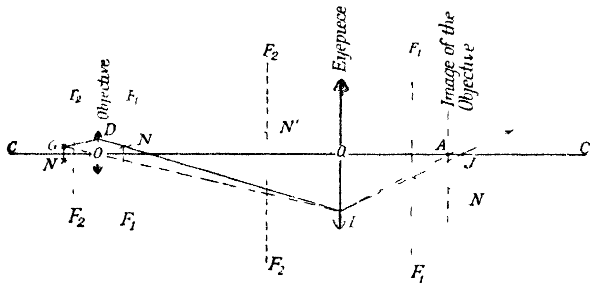


Fig. 25 —Path of the rays in the microscope

Aperture Stop.—The pencil of rays starting from the point G of the object has its angular aperture limited by the mount of the front lens of the objective (or by a stop included in the objective system behind this lens); this mount (or the stop) is therefore the *aperture stop* of the microscope.

Exit Pupil.—We construct the eyepiece image of this aperture stop of the objective by joining OE and determining the direction of the refracted ray EA by means of the subsidiary axis parallel to OE (not drawn in the figure). A is the centre of the image of the aperture stop. All the rays which pass through the objective and fall on the eyepiece also pass through the image of the aperture

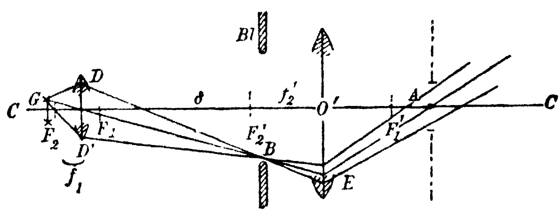


Fig. 26.—Field stop in the microscope

stop; the central rays of all the pencils forming the image must therefore intersect at A. The image at A of the aperture stop is the exit pupil of the microscope. If the observer's eye is to receive as many as possible of the rays passing through the microscope, it must accordingly be in such a position that the exit pupil at A coincides with the pupil of the eye.

Field Stop.—We obtain all the rays which start from the point G of the object and share in the production of the image by tracing out the course of the cone of rays whose vertex is at G and whose base is the objective; this is shown in fig. 26. The diverging cone of rays entering the objective is brought to a point at B. The rays thereafter diverge again, and, as B lies nearer the eyepiece than its first focus, they are still divergent when they leave the eyepiece. The base of the cone of diverging rays is the exit pupil previously indicated.

B, the point to which the cone of rays converges, is the real image of G produced by the objective. In this position a stop Bl is set up, which limits the real image. This stop is a *field stop*, for if the opening of the stop is made small, it is only central rays GB starting from points very near the axis that get through the microscope.

Lateral Magnification.—The real image B (fig. 26) is produced very close to the focal plane F_2'.

If the image B coincided exactly with the focal plane F_2' (fig. 27), the rays emerging from the eyepiece would be parallel; to an eye at A it would appear as if the rays came from an infinite distance, so that no accommodation would be required. The microscope is frequently adjusted in such a way that the virtual image appears at the least distance of clear vision. (According to other experiments, observers more frequently use an accommodation of 1 to $1\frac{1}{2}$ dioptres in looking through an optical instrument (p. 124)). Hence the rays on emerging

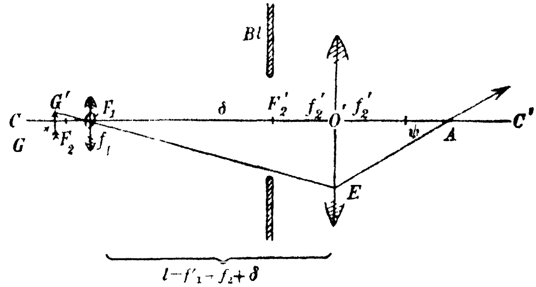


Fig. 27.—The apparent magnitude of the image seen through the microscope

from the eyepiece must appear to diverge from the point which marks the least distance of distinct vision. As compared with the small focal length (only a few millimetres) of the eyepieces used in microscopes, however, the least distance of clear vision (25 cm.) is always to be regarded as comparatively large. Hence B, the point from which the pencil entering the eyepiece diverges, always lies very close to the second focal plane of the eyepiece.

The distance (δ) between the two neighbouring foci within the microscope, the optical tube-length, is always large (it is usually about 16 cm.) compared with the focal lengths of the objective and the eyepiece (which amount to a few millimetres at most). The second focus of the microscope as a whole is the real image of the second focus of the objective. Owing to the length of the tube, however, this latter focus and the aperture stop are both at "great" (and nearly equal) distances from the first focus of the objective (great, that is, in comparison with the focal length of the eyepiece). Their images, i.e. the second focus of the whole microscope system and the exit pupil A, lie very close to one another in the neighbourhood of F_1', the second focus of the eyepiece.

Hence if the eye placed at A sees the image formed by the microscope at the distance d equal to the least distance of distinct vision, —d may without appreciable error be taken as equal to x', the distance of the image produced by the microscope from the second focus of the microscope. But by p. 78 the lateral magnification due to an optical system is given by $m = -x'/f'$, where f' is the second focal length. For a system consisting of two optical systems situated in the same medium,

the second focal length is, by p. 85, given by $f' = -f_1'f_2'/\delta$, where f_1', f_2' are the second focal lengths of the separate systems and δ is the distance between neighbouring foci. Hence, putting $d \approx x'$, $f_1' = f_1$, $f_2' = f_2$, we have

$$m = \frac{x'\delta}{f_1 f_2} = -\frac{\delta d}{f_1 f_2},$$

a result which is due to ABBE.

If we put $d = \frac{1}{4}$ m., reckon δ in metres and the refracting powers $D_1 = 1/f_1$ and $D_2 = 1/f_2$ in dioptres, we have

$$m = -\frac{\delta}{4} \cdot D_1 D_2.$$

The magnification of a microscope is proportional to the least distance of distinct vision of the observer and to the distance between the neighbouring foci of the two lenses, and inversely proportional to the product of the focal lengths.

It follows from the illustrations given above that the image observed by the microscope is *inverted* relative to the object

8. Compound Objectives and Eyepieces.

Telescope Objectives.—In investigating the mode of action of optical instruments, we have hitherto assumed that both objective and eyepiece consist of a single lens of negligible thickness. In actual applications, however, we have always to do with systems consisting of several lenses. For the objective of a telescope, a system consisting of a convergent lens of crown glass and a divergent lens of flint glass is generally used (p. 166). This does away with the coloured margins which appear in the images produced by a simple lens. In addition, it is possible to get rid of spherical aberration for the most intense rays by suitable choice of the radii of curvature of the lenses.

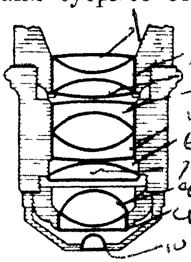


Fig 28.—Microscope objective consisting of ten lenses, apochromatic system

Microscope Objectives.—The objective of a high power microscope has to satisfy particularly stringent requirements in order that the image may exhibit neither colour nor distortion in spite of the fact that it is formed by a pencil of very wide angular aperture (cf. p. 106). Hence microscope objectives always consist of a number of lenses of different kinds of glass. Fig. 28 shows a magnified section through a so-called *apochromatic system* * consisting of ten lenses.

* ABBE applied the term *apochromatic* to objectives in which not only the chromatic aberration along the axis is diminished, but the chromatic difference of the spherical aberration for two colours is got rid of. Here the use of fluorspar (fluorite), first adopted by ABBE in 1886, plays a very important part.

Field-lens and Eye-lens.—The eyepieces of telescopes and microscopes also consist of compound lens systems designed so as to minimize spherical aberration, astigmatism, distortion, and chromatic aberration (Chap. VII, § 6, p. 165). As a rule the eyepiece consists of two lenses; that nearest the object (the *field-lens*) performs a duty similar to that of the condensing lens of the projection lantern, i.e. it is to collect all the pencils of rays which go to make up the image, so that all the light reaches the pupil of the observer's eye. If no other field-stop is present, the mount of the field-lens determines the size of the field of view. The other lens of the eyepiece is the eye-lens, which is specially concerned in the magnification.

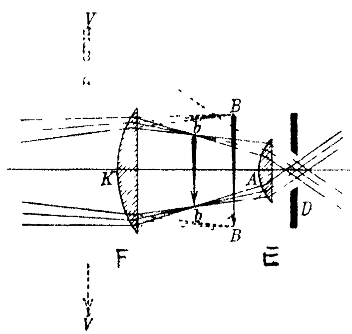


FIG. 29.—Huygens eyepiece with (large) field-lens and (small) eye-lens

The Huygens Eyepiece. — This eyepiece (1703) consists (fig. 29) of two plano-convex lenses with their curved sides nearest the objective and at a distance apart equal to half the sum of the focal lengths of the separate lenses. The field-lens K is so placed that the rays from the objective reach it before they form the real image BB′; hence it displaces this image to *bb*′. The rays proceeding from the latter are then made parallel in the case of the telescope by the smaller eye-lens A; to the observer looking through D, the cover of the eyepiece, the image VV therefore appears at an infinite distance.

If the observer uses a little accommodation, the emergent pencil of rays retains a slight divergence, and the virtual image VV may be, e.g. at the least distance of distinct vision. As the real image *bb*′ is produced in the interior of the eyepiece, this is the only place where cross-wires or another form of graticule may be inserted, a fact attended by certain inconveniences. If f_1' is the focal length of the field-lens, f_2' the focal length of the eye-lens, d the distance between the lenses, and f_2 the focal length of the eyepiece as a whole, we have $f_1' : f_2' : d : f_2 = 2 : 6 : 4 : 3$ approximately.

The Ramsden Eyepiece.—This eyepiece * (fig. 30) consists of two identical plano-convex lenses, with their curved sides facing one another and at a distance apart about equal to the focal length of either. The eyepiece is placed so that the rays from the objective meet before they enter the field-lens; the real image BB′ produced by the objective lies outside the eyepiece. Hence cross-wires or another form of graticule may be attached to a material stop placed in this position, and these will be reproduced along with the object and at the

* Described in 1783 by JESSE RAMSDEN (1735–1800), the son-in-law of JOHN DOLLOND (p. 166) and his successor in business.

same place (p. 137). In the telescope the rays leave D, the cover of the eyepiece, as a parallel pencil; hence to an observer the image VV appears at an infinite distance.

If f_1' is the focal length of the field-lens, f_2' that of the eye-lens, d the distance between the lenses, and f_2 the focal length of the eyepiece as a whole, we have $f_1' : f_2' : d : f_2 = 1 : 1 : 1 : 1$. Here again, then, d is approximately equal to half the sum of the individual focal lengths, as is important if chromatic aberration is to be avoided. (Chap. VII, § 6, p. 165).

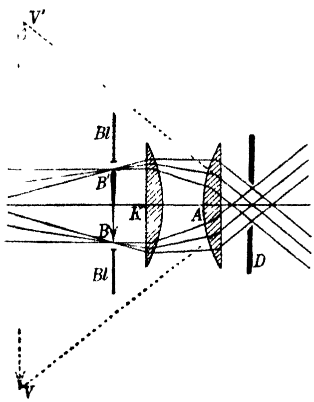


Fig. 30.—Ramsden eyepiece with field lens and eye-lens identical

In the microscope both the Huygens eyepiece and the Ramsden eyepiece produce real images of the objective lens just behind the eye-lens, so that the pupil of the eye may be brought close to them. The image lying at infinity (or at the least distance of clear vision) may then be observed through this exit pupil of the whole microscope as if through a peephole (p. 132). To see the marginal parts of the field of view, the observer must move the head or eye slightly.

More modern types of eyepieces contain a greater number of lenses and remove defects in the image more completely.

The **brightness** of the image produced by an optical instrument depends on the quantity of light which, starting from the object, traverses the stops and lenses of the optical instrument, and finally reaches the eye of the observer. In both the astronomical telescope and the terrestrial telescope, as well as in the microscope, the objective acts as the aperture stop, limiting the rays forming the image. Now as in the telescope all the rays enter at a very small angle to the axis, the brightness of the image depends only on the diameter of the objective and hence is only limited by the practical difficulties of making a very large objective (p. 136) and, in particular, by the high cost of doing so. In the microscope, on the other hand, the diameter of the objective cannot be increased indefinitely, as the objective of a high-power microscope must have a very short focal length and a lens of short focus cannot be given a large diameter. Hence if an object under a microscope is to send as much light as possible into the microscope to form the image, the cone of rays entering the microscope must have a very wide aperture.

A microscopic object is usually placed on a slide consisting of a piece of plate glass bounded by parallel planes, embedded in suitable material, e.g. Canada balsam, and shielded by a thin cover-glass. The position of an object P relative to the front lens O of a microscope objective is shown on a magnified scale in fig. 31. TT is the slide, P the object, CC the layer of material in which the object is embedded, DD the cover-glass, and O the semicircular front lens of the objective, which is separated from the cover-glass by a thin layer of air. Of the light

rays starting from P only a pencil whose aperture ψ is less than twice the critical angle of total reflection can reach the layer of air and hence the objective; rays which lie outside this angle, such as R, are totally reflected.

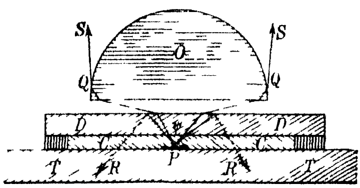


Fig 31.—Microscopic object and cover glass (dry immersion)

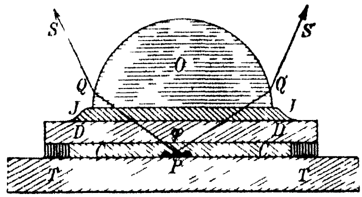


Fig 32 —Microscopic object and cover-glass (homogeneous immersion)

Immersion Systems.—To increase the aperture (cf. p. 57) a suitable liquid, e.g. cinnamon oil,* is placed in the space between the cover glass and the lower surface of the front lens. A microscope objective system which has been calculated for use with a layer of oil in this way is called an **oil-immersion objective.** If the oil has the same refractive index as the cover-glass and the front lens of the microscope, we speak of **homogeneous immersion.** In contradistinction to this, systems used without a liquid are referred to as **dry immersion systems.** The way in which a homogeneous oil immersion works is clear from fig. 32. The aperture ϕ of the cone of rays starting from P is only limited by the fact that the object P cannot lie immediately on the lower surface of the front lens. The angle ϕ may reach nearly 180°, as the rays travel in straight lines from P till they leave the

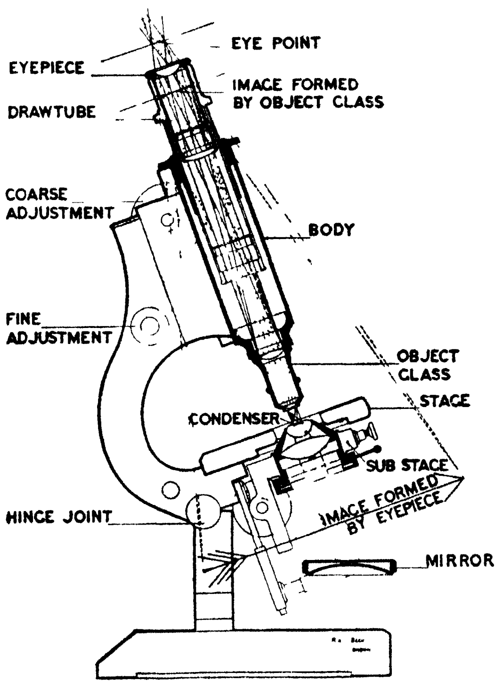


Fig. 33 —Section through a microscope (R. and J. Beck, Ltd.)

* Due to AMICI (1840).

front lens O. With *water-immersion* systems such large apertures cannot be attained (fig. 35, p. 57). Homogeneous immersion permits the use of front lenses consisting of more than a hemisphere (fig. 26, p. 105).

Resolving Power.—In addition to the brightness, the *resolving power* (§ 4, p. 121) of the microscope is considerably increased by the use of an immersion system, an account of the increase in aperture. This is due to the fact that diffraction effects arising from the undulatory nature of light are thereby minimized.

The brightness of the microscope image naturally depends essentially on the illumination of the object also. As a microscopic object is very small and the rays must diverge strongly as they enter the objective in a cone of wide aperture, the rays illuminating the object also act most effectively if they are strongly convergent when they reach the object. Following out a suggestion of ABBE's, therefore, a system of lenses is placed beneath the stage, to make the rays coming from the mirror below convergent.

Fig. 33 shows a section of a type of microscope in common use.

9. Effects of Diffraction on the Production of Images.

Diffraction Discs.—Hitherto we have based our discussion of the action of optical instruments on geometrical or ray optics, and we have made ourselves acquainted with the conditions under which an object-point gives rise to a single image-point, i.e. the conditions under which a proper "image" is formed. But even when all the errors previously mentioned have been avoided, the image of the object-point is not sharp. The reason for this lies in the undulatory nature of light.

Any lens holder acts as a slit, the diffracting effects of which were discussed * in Chap. I, § 6 (p. 16); see also Vol. II, p. 244. If parallel rays fall on a lens (say from a star), the effect of the lens and its mount is shown diagrammatically in fig. 32, p. 199. As the lens K may be regarded as a diffracting body G (p. 199), we obtain the simplified diagram of fig. 34. According to geometrical optics, the convex lens K brings all the parallel rays L to a focus at O. If, however, we consider the diffraction at the opening, we find that at the point A on one side of O light waves of different phases meet and produce maxima and minima of intensity. Fig. 35 shows how the intensity diminishes from the centre (the actual "image") gradually, but with the occurrence of secondary maxima. The complete configuration is to be obtained by imagining the figure rotated about the axis of ordinates.

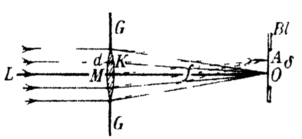


Fig. 34.—Diffraction patterns due to the mount of the objective

* For the quantitative discussion, see Chap. VIII, § 8, p. 198.

If the radius of the mount of the lens is d, the focal length f, and the wave-length of the light λ, the radius of the first dark diffraction ring is given * by $\delta = 0{\cdot}61\dfrac{f\lambda}{d} = 0{\cdot}61\dfrac{\lambda}{d/f}$. Thus in place of the point which we should expect from geometrical optics, there appears a small disc, the so-called "diffraction disc", whose diameter is given by $1{\cdot}22\dfrac{\lambda}{d/f}$ and whose brightness decreases from the centre outwards. It is also surrounded by concentric rings or by various other figures varying in shape with the shape of the opening (fig. 36, Plate XII). The latter figures are usually so feeble that they need not be taken into account except in special circumstances. If the light used is parallel and the aperture of the incident pencil is always less than that required for the first elementary Fresnel zone, the centre of the diffraction figure is a bright spot of light, the diffraction disc (cf. also Vol. II, fig. 28, left, p. 245) or actual " image " of the star.

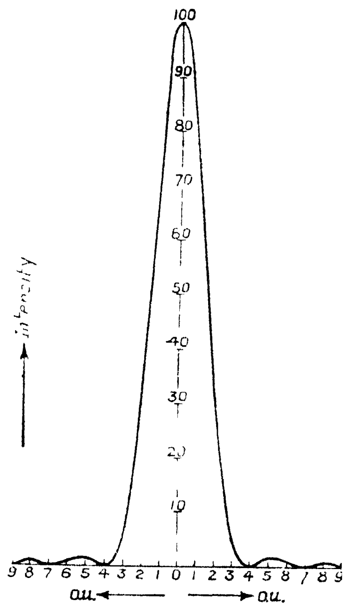


Fig. 35.—Distribution of brightness in the "image" of a point: o.u. = optical units, 1 o.u. = $f\lambda/2\pi d$ (From *Handwörterbuch der Naturwissenschaften*, Vol. I (G. Fischer, Jena))

If, as in the microscope,† we are concerned with the formation of an image by a system of maximum numerical aperture A, we may also write the above formula for the radius of the diffraction disc in the form $\delta^a = 0{\cdot}61\lambda/A$, if we think of the radius projected back into the object, i.e. state its magnitude relative to the details of the object.

Depth of Focus.—The image of a point of the object is not sharp in the depth direction either, that is, the interference phenomenon extends along the axis also. The light energy is distributed uniformly, so far as we can tell, over a region of a certain depth. This so-called depth of focus (BEREK) is given by $T = 1\lambda\mu/A^2$, where μ is the refractive index of the object.

* Except for the factor 0·61 the equation is identical with the equation for diffraction at a slit (p. 16); the factor 0·61 follows from the strict theory of diffraction of parallel rays at circular openings.

† The importance of diffraction effects in the formation of images by the microscope was first recognized by E. ABBE (p. 53); this formed the starting-point of enormous developments in the instrument.

Fig. 37 gives the interference image of a luminous point as it would appear if perfectly reproduced (according to geometrical optics), the light used being monochromatic. AA is the optic axis, BB the plane of the "image". The whole configuration is obtained by

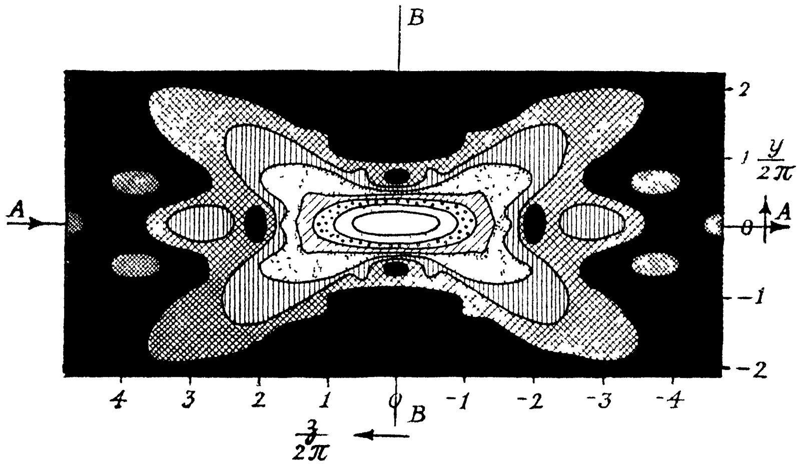


Fig. 37.—Image of an indefinitely small element of surface radiating light in accordance with Lambert's law and reproduced without aberration according to geometrical optics (after Berek).

The distribution of light is projected back into the object-space, so that the figure gives the relationships in due proportion as they would appear if the optical instrument were moved along its axis. If the brightness in the very centre of the figure is taken as 100, the various shadings have the following meanings: in the completely black region the brightness is less than 1, in the succeeding regions it is less than 2½, 5, 10, 25, 50, 75, 100 respectively. AA is the optic axis, BB the plane of the image. The scale is graduated in optical units (see legend of fig. 35). (From *Zeitschrift für Physik*, Vol. 40 (Springer, Berlin).)

imagining the figure rotated about AA. Owing to the defects of optical systems still greater deviations from an ideal "image" will occur.

Resolving Power.—By this we mean the power of an instrument to reproduce two separate but neighbouring points of the object as separate points (p. 121). If the two points are so situated that their diffraction discs actually touch, they can certainly still be recognized as separate points, as the brightness of the discs decreases outwards. For the latter reason the images may still be recognized as separate when the diffraction discs partly overlap. The resolving power is therefore determined by $\delta^a = 1{\cdot}22\lambda\kappa/A$, where κ denotes a factor which expresses the physiological resolving power for differences of brightness and may on the average be taken as 0·5.

Thus, for example, if two stars observed by a telescope are very close together, the telescope will not be capable of resolving the

"double star" unless the distance between the centres of the two diffraction discs exceeds $0{\cdot}61\frac{\lambda}{d/f}$. If the angular separation of the two stars is ϕ, the distance between their images, that is to say, the distance between the corresponding diffraction discs, is $f\tan\phi$, which is approximately equal to $f\phi$. Hence we must have $0{\cdot}61\frac{\lambda}{d/f} \leqslant f\phi$, or $0{\cdot}61\frac{\lambda}{d} \leqslant \phi$. For this reason it is essential that d, the radius of the telescope objective, should be sufficiently large. Hence it follows that the objective of a telescope which has a long focal length must also have a large diameter, in order that the maximum resolving power may be attained. Increasing the size of the eyepiece cannot increase the resolving power of a telescope; by increasing it too much we merely obtain an "empty magnification" which reveals no further details.

From fig. 35 it follows immediately that the "image" is sharper the greater the diameter of the objective ($2d$). Hence the brightness of the diffraction pattern is proportional to the *fourth* power of the diameter of the objective; for if the diameter of the objective is doubled, the quantity of energy entering is quadrupled and the area of the diffraction pattern is at the same time divided by four. Thus the brightness at a point of the "image" is sixteen times as great as before. This is another reason why the diameters of objectives and mirrors for astronomical purposes are made so large (p. 136).

A telescope with an objective 20 cm. in diameter is capable of separating two stars whose angular distance apart is 0·6″; the resolving power relative to angular distance is proportional to the diameter of the objective, but is independent of its focal length and hence of the magnifying power of the telescope. A photographic plate is in general most sensitive to rays of shorter wave-length than those to which the eye is most sensitive (p. 30). Hence a higher resolving power is attained by photographing a double star than by observing it with the eye.

Similar considerations apply to the microscope. As $1{\cdot}22/A$ amounts in the most favourable case (for an immersion system) to 0·9, it follows from the expression for δ^a that points of the object can still be recognized separately if their actual distance apart is about $\lambda/2$. As the wave-lengths of violet and red light respectively are nearly in the proportion of 1 : 2, the resolving power can be doubled by using violet light instead of red light. A further advance may be secured by the use of ultra-violet light (usually about 280 $m\mu$) and photography (A. Köhler). As the light used is almost monochromatic, *M. von Rohr's monochromats*, which are not corrected for chromatic aberration, may be used.

If the object under examination has a periodic structure (e.g. a diffraction grating or diatoms) peculiar interference phenomena arise, which formed the starting-point of Abbe's investigations into the formation of an image by the microscope; these, however, apply only to periodic structures. Complicated interference may occur, so that in certain circumstances no proper image plane may exist at all.

Thus in order to make fine details recognizable under the microscope, the point is not to increase the magnification indefinitely, but in the first place to choose the aperture of the objective in such a way as to give the desired resolving power. The magnification of the eyepiece is then to be chosen in such a way that everything that the objective is capable of resolving in virtue of its aperture can be readily recognized by the eye, i.e. subtends a sufficiently large angle at it. This is the case when the magnification of the eyepiece m_{EP} is selected so as to lie between 500 and 1000 times A/m_{OB}. If the magnification of the eyepiece is less than this the power of the objective is not fully utilized, while if it is greater, an " empty " magnification (see above) results.

Condensers.—These are used to illuminate the object under investigation as brightly as possible and with a cone of rays as wide as possible, so that the object may be regarded as self-luminous, owing to the large number of incoherent rays diffracted simultaneously (LORD RAYLEIGH). Fig. 38 shows the essential parts of a condenser and the path of the rays through the condenser and the microscope objective when the condenser stop is fully open.

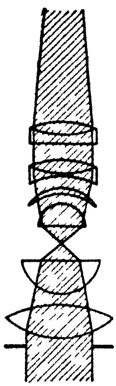


Fig. 38.—Bright-ground illumination with condenser

Dark-ground Illumination.—If a thin pencil of sunlight crosses the room, we see minute particles of dust sparkling in it, which are at other times invisible owing to their smallness. The reason is that the sunlight is diffracted by the particles and their apparent magnitude is considerably increased by the formation of diffraction discs. If the particles are very close together, as e.g. in tobacco smoke, we can no longer detect them separately with the naked eye, but we merely see them as a whole as a diffusely reflecting pencil of light. If, however, we examine this pencil with a microscope of high resolving power, the bodies, magnified by the diffraction discs, become separately visible.

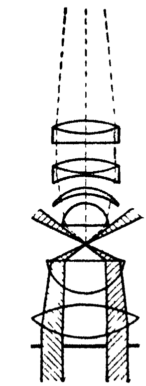
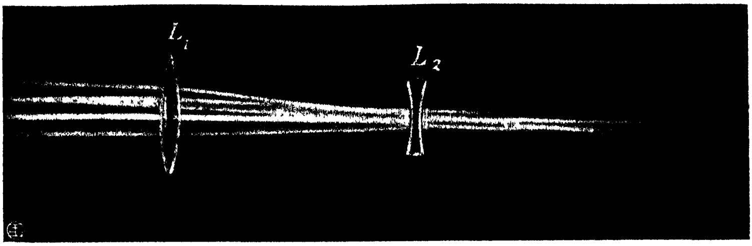


Fig. 39.—Dark-ground illumination by means of central stop

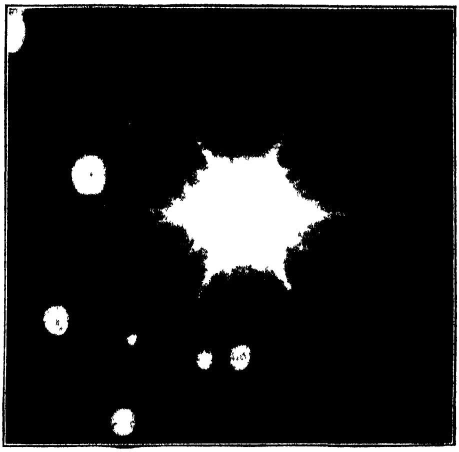
For this purpose we may stop down the condenser as shown in fig. 39 or use a condenser of special construction (fig. 40, Plate XII).

For detecting very small particles the following apparatus, which is known as an **ultramicroscope** and was first described in 1903 by SIEDENTOPF and ZSIGMONDY, is very frequently used.

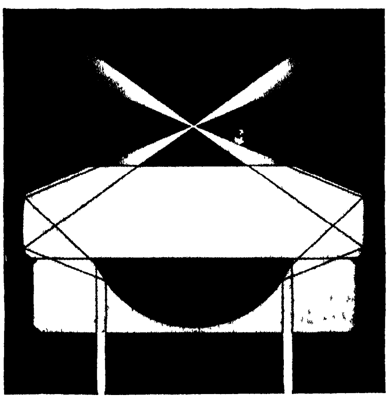
In fig. 41 let G be the object under examination, e.g. a liquid with small particles in suspension, contained in a rectangular glass dish. L is a very narrow horizontal slit which is illuminated from the left by a bright source of light, so that a narrow pencil of rays gets through; this pencil is made to converge slightly by the convex lens B. A very thin horizontal layer of the liquid G is thereby illuminated intensely.



Ch VI, Fig 35 Path of the rays in the Galilean telescope (p 127)



Ch VI, Fig 36 Diffraction pattern due to a star (ι Orionis) photographed by a reflecting telescope (p 145)



Ch VI, Fig. 40 — Cardioid condenser (after Siedentopf) (Carl Zeiss, Jena)

The small particles suspended in the liquid diffract the light, and part of the light travels vertically upwards into the objective O of the microscope M. These diffracted rays may be observed against the dark field of view (fig. 42). In the part of the body traversed by the beam of light there appear small isolated points of light, which indicate that at the corresponding part of the body there is a particle which is

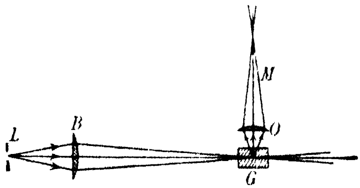


Fig 41 Arrangement for dark-ground illumination in the ultra-microscope

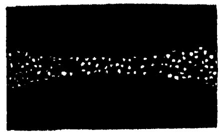


Fig. 42.—Ultramicroscopic image

diffracting the light. The bright points observed are not actual images of the bodies, but interference patterns, which are merely a sign of the presence of the particles but do not directly indicate their shape. Under certain conditions, however, the nature and form of the particles may be inferred from the form of the interference pattern. This so-called **interference microscopy,** however, is only in its earliest beginnings.

By means of the ultramicroscope it is possible to detect particles as small as 4 $m\mu$. The limits of visibility depend on the fact that smaller particles, even when very intensely illuminated by sunlight, are incapable of sending a sufficient quantity of diffracted light into the microscope.

CHAPTER VII

The Influence of Wave-length on Refraction Phenomena

1. The Decomposition of White Light.

The Spectrum.—We have already found when discussing diffraction phenomena (Chap. I, § 6. p. 11) that colours can arise from white light. In fact if we look at a distant source of light at night through the half-closed eyelids we see coloured diffraction fringes, which are produced by the criss-cross of the eyelashes. On p. 19 we traced back the occurrence of coloured interference bands or fringes to the differing wave-lengths of the light used, and found how to measure the latter. As we shall show in what follows, it is possible by means of refraction to carry out the separation of the wave-lengths very conveniently and to exhibit the separate wave-lengths in high intensities. We shall therefore begin by using this method to investigate the phenomena of colour sensation which are most directly perceptible to the senses.

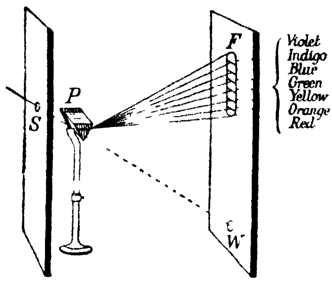


Fig. 1.—Decomposition of white light

If a pencil of parallel rays of sunlight limited by a circular stop falls on a white screen in a dark room, it gives rise to a round white spot on the screen. If in the path of the rays we interpose a glass prism with the refracting edge downwards (fig. 1), so that the path of the rays through it is fairly symmetrical, the beam is not only deviated, as we know, but at the same time it is spread out like a fan; the white spot on the screen vanishes and a coloured band F appears higher up on the screen, its upper end being violet and its lower end red. The violet part of the fan is accordingly the most deviated by the prism, and the red the least deviated. Between the two outermost colours a large number of different colours are included; their number cannot be stated exactly owing to the gradual nature of the transitions between them. It is, however, customary (following NEWTON) to select the seven typical colours red, orange, yellow,

green, blue, indigo, violet. These are known as the colours of the rainbow. The whole coloured band is called the **spectrum.***

The emphasis on the "seven" fundamental colours of the spectrum arose from NEWTON'S attempts to compare the spectrum with the octave in sound. Experiments have shown that about 160 different colour sensations can be distinguished from one another in the spectrum.

If we blow smoke into the path of the deviated fan of rays in order to make their course through the air visible, we find that just beyond the prism only the outer margins of the fan are coloured, the centre being pure white, but that the colours appear more clearly the greater the distance from the prism. If we were to attempt to separate the colours completely in this way, the spectrum would soon become too feeble.

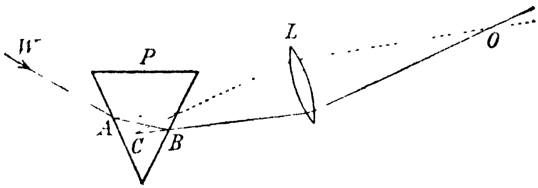


Fig. 2.—Recombination of the coloured band to form white light

Dispersion.—As was first demonstrated by NEWTON (1704), white sunlight is not a pure light but a combination of the many-coloured rays of the spectrum. The prism causes **dispersion** of the compound light, because the differently coloured constituents are not deviated to the same extent; the refractive index of glass is *less* for *red* light than for *violet* light.

This also explains the fact that just beyond the prism the margins of the beam only are coloured, the centre remaining white. For at the centre of the beam the parts are not yet completely separated: the different parts of the different rays overlap and thus produce white light.

By recombining the coloured constituents, white may again be produced.

If for this purpose we let the coloured rays fall on a series of small mirrors which are directed in such a way that the reflected coloured rays illuminate the same spot on a white screen simultaneously, this spot will appear white. The rays may also be recombined by means of a large concave mirror or a large convex lens. The latter experiment, illustrated in fig. 2, is particularly instructive; it shows that the point O where the rays meet is actually white, but that beyond this point the rays diverge again, and in an order which is the reverse of that which they had before they reached the point O.

The Pure Spectrum.—A pure spectrum may be produced, without

* Lat., *spectrum*, appearance.

lowering the intensity too greatly, by using the following arrangement due to FRAUNHOFER * (fig. 3). The essential point is the replacement of the round hole by a slit parallel to the refracting edge of the prism. Through the narrow slit S, which is illuminated as brightly as possible by sunlight from the left, light rays fall on the convex lens L, which causes them to converge, forming a real image of the slit at the point B. The prism is set up immediately behind the lens in such a way that its refracting edge is parallel to the slit. Every pencil of rays which would be brought to a focus at the point B if the prism were not there is subject to deviation and dispersion by the prism. Only the limiting rays (violet and red) are drawn in the figure. All the red rays now form an image R, all the violet rays an image V, and all the rays of other colours images lying between R and V. Thus a spectrum is formed which can be caught on a screen and which consists, owing to its method of production, of *a succession of separate coloured images of the slit.*

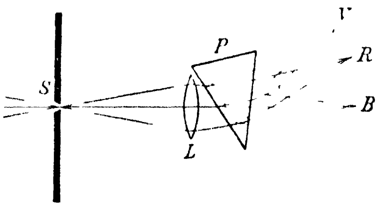


Fig 3 —Fraunhofer's apparatus

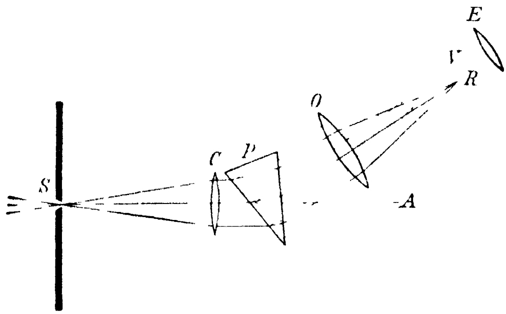


Fig. 4 —Another arrangement due to Fraunhofer

If we replace the rectilinear slit S by a slit of any other form, the individual parts of the spectrum each take on the new shape of the slit.

The arrangement shown in fig. 4, also due to FRAUNHOFER, is still more complete. In the path of the rays from the slit S is placed the convex lens (collimating lens) C, in such a way that the rays leave it parallel to one another, i.e. the slit S lies in the focal plane of the lens C. Immediately behind C the prism is set up so that the parallel rays traverse the prism symmetrically, i.e. at minimum deviation. After the rays have been deviated and dispersed by the prism, they fall on another convex lens O, which causes each system of parallel rays to form

* JOSEPH VON FRAUNHOFER (1787–1826), the pioneer of scientific instrument manufacture in Germany, was professor at Munich from 1823 till his death; he did important work in optics and is specially famed for his discovery of the spectral lines, his work on diffraction, and his improvements in telescopes.

an image of the slit in its focal plane. Then the individual images of the slit again form a *pure* spectrum. The advantage of this arrangement lies in the fact that all rays of the same colour traverse the prism under the same conditions, i.e. are subject to the same amounts of deviation and dispersion, whereas with the arrangement of fig. 3 the individual rays leave the prism convergent, i.e. at different angles.

In carrying out the above experiments we may catch the spectrum on a white screen, i.e. it may be exhibited objectively. We may also set up a convex lens E behind the spectrum VR and observe the spectrum subjectively through it as if through a magnifying-glass.

The Spectroscope or Spectrometer. The two lenses O and E may be regarded as forming the two constituent lenses, the objective and the eyepiece, of a simple astronomical telescope. If they are united to form a single piece of apparatus, we have the *spectroscope*, which was first constructed by KIRCHHOFF* and BUNSEN†, and which is shown in fig. 5.

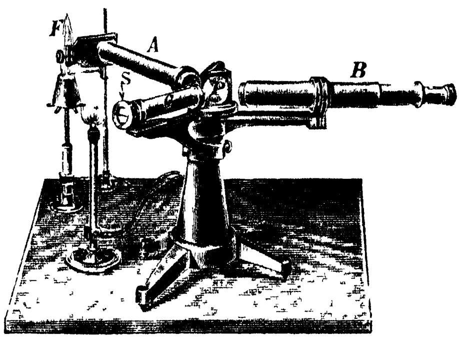


Fig. 5.—Bunsen and Kirchhoff's spectroscope

The triangular prism P is placed on a firm support or table. One face of the prism is directed towards the collimator tube A, which at the end next the prism carries a convex lens (the collimating lens) and at the other end, at a distance equal to the focal length of the lens, a slit which can have its width adjusted by a micrometer screw. The other face of the prism is directed towards the telescope B, which is focused for infinity. The light passing through the slit is rendered parallel by the collimating lens; it is then deviated and dispersed by the prism and reaches the eye of the observer by way of the telescope B. Usually there is a third tube, the scale tube C, as well, carrying a convex lens at the end next the prism and a scale S at the other end. The scale tube is fixed in such a way that the rays emitted by the illuminated scale are reflected by the front surface of the prism and then enter the telescope B. The observer accordingly sees the spectrum and the scale one above the other and is thus in a position to measure the positions of the various parts of the spectrum.

The Spectrograph.—Spectra are usually recorded photographically. The telescope is then replaced by a camera (fig. 6, Plate XIII). To decompose the light more effectively, several prisms are often used. If a slit is placed in the focal plane of the spectrograph instead of the photographic plate, it is possible to select a single colour or a restricted region of the spectrum from the compound light emitted by a source of light (*monochromator*). This apparatus is usually so made that different spectral regions can be selected by means of a fixed slit, by automatically rotating the prisms so that the wave-length required traverses them at minimum deviation. For the *resolving power* of the prism spectrograph see p. 211.

* R. KIRCHHOFF (1824–1887), Professor of Physics in Heidelberg from 1854 and in Berlin from 1878. † See p. 158.

2. Colour Mixtures: Complementary Colours.

Recombining the Colours of the Spectrum.—When all the parts of the spectral fan of rays into which white light is split up are brought together again, white light results. If, on the other hand, some parts

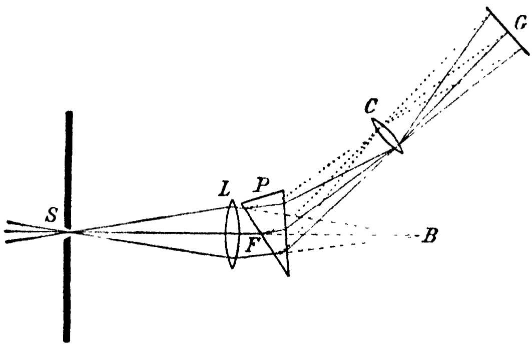


Fig. 7.—Recombination of a pure spectrum to form white light

are previously removed by the insertion of screens or deviated by a small prism, the rest of the fan of rays unites to give a special mixture of colours.

Fig. 7 gives a diagrammatic representation of the apparatus by which a pure spectrum may be recombined to form white. In the absence of the prism P the lens L would unite the rays entering through the slit S to form a real image of

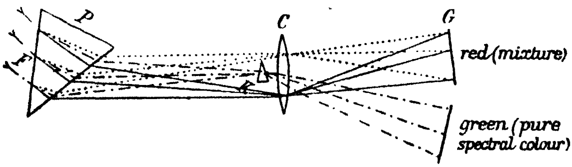
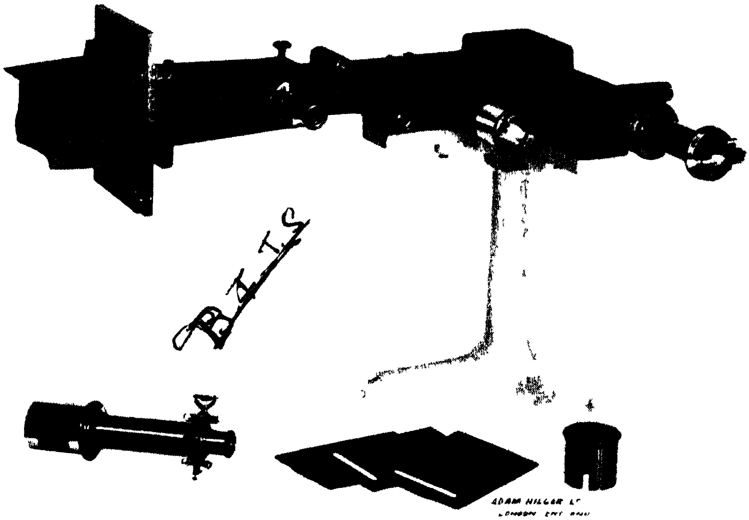


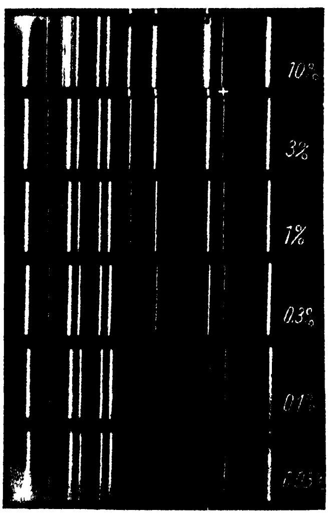
Fig. 8.—When the pure spectral green is deviated to one side the remainder of the spectrum forms a red mixture

the slit at B. The prism deviates and disperses each separate ray, giving rise to the pure spectrum at C. If we now insert at C a convex lens whose focal length is approximately half the distance of C from the front face F of the prism P, the lens will reunite the rays of various colours diverging from every point of the face F at a point G behind the lens, at a distance which is again equal to twice the focal length of C; at G there results a real inverted colourless image of the surface F. As the rays starting from F diverge only in the plane of the principal section of the prism, but not in planes parallel to its refracting edge, the lens C may be a cylindrical lens with its curved surfaces perpendicular to the plane of the figure.

Colour Mixtures.—If we now insert a thin prism K with a small refracting angle immediately in front of C (fig. 8 represents the part



Ch VII, Fig 6 Spectrograph (Adam Hilger Ltd)



Ch VII, Fig 10 Scale of comparison spectra for deducing the composition of a mixture of tin and a small percentage of cadmium The tin line (+) 3656 Å is compared with the three lines of the cadmium triplet (|), 3615 Å, 3613 Å, 3611 Å (which are not separated), 3468 Å 3466 Å (not separated) and 3404 Å (p 158)

(From Gerlach-Schweitzer, *Die chemische Emissionsspektralanalyse* (Voss, Leipzig English translation, *Foundations and Methods of Chemical Analysis by the Emission Spectrum*, published by Hilger))

to the right of the prism in fig. 7 on a larger scale), we may deviate e.g. the green rays to one side. In addition to the image of the face F, which was originally white, but now appears coloured, we then have a green image of the face F due to the green rays alone. The image of the face F, which originally was white, now appears to have the *mixed colour* red resulting from the rest of the spectrum. As a rule such colour mixtures are subjectively indistinguishable from the pure spectral colour. The only colour mixtures not represented in the spectrum are the shades of *purple* obtained by mixing red and violet.

Complementary Colours.—Any portion of the rays may be removed in the manner just described, by moving the thin prism K along the pure spectrum. The results obtained are as follows:

Deviated part	red	orange	yellow	green	blue	indigo	violet
Remaining colour mixture	bluish-green	blue	indigo	purple	orange	yellow	greenish-yellow

If we recombine the deviated portion and the remaining colour mixture, we of course obtain white again. The pair of colours in each vertical column of the table when combined give white; hence they are known as pairs of *complementary colours*.

Although subjectively they have the same effect, the pure spectral colours in the upper row of the table exhibit a physical behaviour differing from that of the colour mixtures in the lower row. For example, if the screen to catch the image is set up not at G but farther away, decomposition into the individual constituents again takes place.

It is also possible to combine individual constituents of the spectrum in pairs to form white. These are then pure complementary colours. They occur in pairs just as in the above table.

The white arising from the combination of *two pure* complementary colours alone, however, is not physically identical with the white of sunlight; in the former case the impression of "white" is solely due to a peculiarity of the eye, for decomposition by a prism or spectroscope again gives rise to the two original colours only and not to the complete spectrum.

Note.—We may mention here in advance that except for the region between 500 and 560 $m\mu$ (compare Vol. III, p. 646) the wave-lengths λ_1, λ_2 of these two pure colours satisfy the formula $(\lambda_1 - 559)(498 - \lambda_2) = 424$; from this it is possible to calculate the wave-length of the pure colour complementary to a given pure colour.

Thus a colour may arise in a great variety of ways. The eye is not capable of detecting this variety of origins; it cannot analyse colours as a practised ear can analyse sounds. For this reason the spectroscope is one of the most important of optical instruments, as it furnishes a reliable means of ascertaining the composition of various lights.

Newton used the **colour top** * to combine or mix separate colours. If a top is painted out in sectors with the colours of the spectrum and rapidly revolved, we obtain the impression of dirty white or grey. We can no longer distinguish the separate colours, but receive the same impression as if the colours were mixed (additive mixing). That the white produced is not pure white is partly due to the fact that it is impossible to paint the top with all the colours in the exact shade and intensity with which they occur in the spectrum, and partly to the fact that the colour top does not emit so much light as if its surface were pure white. The mixing of complementary colours may also be carried out with the colour top.

Body Colours or Pigments.—These are essentially different from the physical concept of a spectral colour. A red piece of cloth only appears red to the eye because it chiefly emits red light. As, however, the cloth is not itself a source of light, the light it reflects must come from a source whose light includes red light. Hence the cloth appears red both in daylight and in ordinary lamplight. If, however, we bring the red cloth into the objective spectrum of sunlight, it only appears *red* when it is in the *red* part of the spectrum; if it really reflects red light only, it will appear *black* when in other parts of the spectrum. Usually, however, the apparently red cloth is not really monochromatic, but reflects rays other than red, although more feebly. If a larger piece of coloured cloth is held so as to catch the whole spectrum, we see at once which parts of the spectrum are absorbed by the cloth and which are re-radiated. Thus we see that body colours which appear identical to the eye often radiate or absorb different parts of the spectrum, often quite narrow regions, to different extents.

If we hold a piece of coloured glass in the path of the rays traversing a spectroscope, or anywhere in the path of the rays in fig. 4, part of the spectrum vanishes more or less entirely; the coloured glass absorbs part of the spectrum and lets through the rest. The mixture of spectral colours that gets through gives the eye the impression of the colour of the glass. The spectrum as altered by the absorption of light by the coloured body is called the **absorption spectrum** of the body. Many substances have a highly characteristic absorption spectrum, which may be used as a test of their presence. On the coloured plate is shown the absorption spectrum of neodymium as exhibited by most of its compounds.

We accordingly have the following experimental facts: the body colour of a pigment depends essentially on the region of the spectrum absorbed by the body when light passes through it. Coloured glass, dyed wool, &c., appear red if the bluish-green which is complementary to red is more or less strongly absorbed from the incident white light; similarly, absorption of the yellow gives rise to a blue body colour, and conversely. In reflected light the colours of pigments often appear different, e.g. fuchsine crystals and dried red ink have a green lustre. This behaviour is due to the fact that the regions which are strongly absorbed are also strongly reflected.

Mixtures of Pigments.—Mixing blue and yellow pigments (e.g. oil-colours or water-colours) usually gives rise to a *green* pigment, not the *white* colour which one would expect from the behaviour of the spectral colours. The reason for this discrepancy is as follows: the yellow pigments keep back almost all the spectral colours belonging to the blue end of the spectrum, whereas they transmit or reflect the rays belonging to the red end of the spectrum (up to green). The blue pigments absorb all the rays belonging to the red end of the spectrum, whereas they transmit or reflect the rays belonging to the blue end of the spectrum (usually beginning with green). Then if a mixture of pigments contains these two pigments close beside and on top of one another, a particle of blue pigment transmits or reflects a mixture of green, blue, and violet on to a neighbouring particle of yellow

* The experiment was actually known to Claudius Ptolemæus (A.D. 150).

pigment; the latter, however, absorbs the blue and violet rays, so that the green rays alone traverse the mixture unaltered (fig. 9). The green colour of the mixture of pigments is deeper in tint the more intimate the mixture of the pigments, for it is only when the pigments are thoroughly mixed that the incident rays meet the two kinds of pigments successively and repeatedly, so as to cause all the rays to be completely absorbed except the green. The green colour, however, is never so brilliant as if a pigment were used which looks green itself; for the yellow and blue pigments in the mixture always absorb part of the green light as well, whereas in a simple green pigment the green part of the incident white light may scarcely be enfeebled at all.

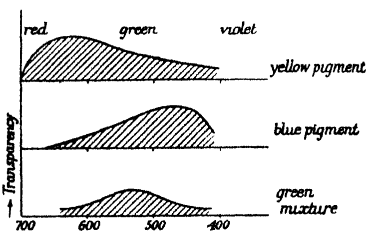


Fig. 9.—Transparency of pigments and mixtures of pigments (the shaded areas denote colours let through)

The colours which arise from repeated absorption of the rays by two or more pigments are called **difference colours** (subtractive mixture of colours). The combination of several spectral colours gives the **summation colour** (additive mixture of colours).

*The difference colour of the two pigments blue and yellow is in general green:** *the summation colour of the two spectral colours blue and yellow is white.*

3. Spectrum Analysis.

Colour and Wave-length.†—By means of a slit placed in the plane of the spectral image, we can select a narrow region of the spectrum produced by a prism and determine its wave-length by means of interference experiments (Chap. I, § 6, p. 11 and p. 208). It is found that when the prism used is of glass or other transparent colourless material the wave-lengths follow one another continuously in such a way that *the rays of longer wave-length are less refracted and the rays of shorter wave-length more refracted.* On p. 646, Vol. III, we have already given a table of the colour sensations and the corresponding regions of wave-length. The connexion becomes still clearer if, as is shown in the central figure of the coloured plate facing p. 158, the interference bands are sketched directly on the spectrum (cf. p. 177). The breadth of the bands is greatest in the red and steadily decreases towards the violet. If a graduated scale is included in the spectroscope (p. 153) it is therefore possible to associate each of its graduations with a definite wave-length. In practice we do not use the interference bands but calibrate the scale with light of known wave-length, i.e. light whose wave-length has already been determined accurately by interference experiments.

Line Spectra.—The method just mentioned is possible in virtue of

* It is not necessarily always so, as the transmission curves may under certain conditions be quite different from those in fig. 9.

† Wave-lengths are usually stated in millimicrons ($m\mu$) or in Ångström units (Vol. I, p. 6). 1 $m\mu$ = 10^{-6} mm. = 10^{-7} cm.; 1 Å. = 10^{-8} cm. (Vol. III, fig. 47, p. 646).

the fact that sources of light exist which emit only light of certain definite wave-lengths. For example, if a sodium salt is heated in a bunsen flame, it gives the latter a bright yellow colour. This yellow light is found to consist almost exclusively of rays of wave-length 589 $m\mu$ (see also the coloured plate). Investigation of this radiation with more powerful apparatus shows that the line visible in the spectroscope, known as the D line, consists of two lines very close together, with wave-lengths $D_1 = 589{\cdot}5932$ and $D_2 = 588{\cdot}9965$ $m\mu$ (in dry air at 15° C. and a pressure of 760 mm.). Spectra characteristic of the various elements are also obtained when electric discharges are passed through gases (Vol. III, Chap. X, § 3, pp. 326–349) and when spark and arc discharges are made to pass between electrodes of various metals. These spectra consist of a smaller or greater number of lines (line spectra; see the coloured plate), which represent the monochromatic images of the slit in the lights of the wave-lengths in question.

Bunsen * and Kirchhoff (1859) are justly renowned for their discovery of the fundamental fact that:

Under given conditions each element emits a perfectly definite spectrum which is characteristic of that element only.

These two scientists were thus not only the founders of modern atomic physics, but by the application of their results to the radiation of the stars they laid the foundations of our present knowledge of the material structure of the universe. For if the lines of an element appear in a spectrum, it may be concluded with certainty that the element is present in the luminous source of light.† This method of investigation is known as **spectrum analysis.** By making the material luminous in some way, e.g. by vaporizing it in a flame or by using it as electrode for an electric discharge, its chemical constitution can immediately be deduced from the spectrum. Under certain circumstances it is even possible to deduce the quantitative composition of the sample from the intensity of the lines, a method which has recently come into practical prominence owing to its elegance and the minute quantities of material required (fig. 10, Plate XIII).

The first few rows of figures in the coloured plate show the spectra obtained when salts of the alkalis or alkaline earths are vaporized in the bunsen flame on a platinum wire. Below these some spectra of discharges through gases are given. The method is so sensitive that quantities as low as $3 \cdot 10^{-7}$ mg. of sodium and $1 \cdot 10^{-7}$ mg. of lithium can be detected. Fig. 10 shows a series of spectra of alloys of tin and cadmium in which the percentage of cadmium decreases steadily downwards. By comparing the spectrum of a similar alloy of unknown quantitative composition with these, its cadmium content may be determined.

* R. W. Bunsen (1811–1899), born in Göttingen, latterly a professor at Heidelberg, was remarkable both as a man and a scientist and is specially noteworthy for the way in which he extended physical methods to chemical problems.

† [See, however, the paragraph in small type on the next page.]

The spectrum of an incandescent solid exhibits all the lines of the visible region of the spectrum, or in other words, it is continuous.*

We see this immediately if we look at the spectrum of an incandescent lamp or a candle flame. The light of the latter is chiefly due to incandescent particles of soot. For further details see the subject of temperature radiation (Vol. V).

Fraunhofer's Lines.—The spectrum of the sun or of a star is continuous also, but certain definite wave-lengths are wanting, so that the spectrum is interrupted by dark bands.

The dark lines in the sun's spectrum were first studied by FRAUNHOFER in 1814. The strongest of these so-called Fraunhofer lines are denoted by letters of the alphabet. Fig. 11 (Plate XIV) reproduces a drawing of these lines by FRAUNHOFER. The wave-lengths of the most important are given in Table II. As will be shown in more detail in Vol. V, the dark lines are to be interpreted as absorption lines, the

TABLE II.—WAVE-LENGTHS OF SOME OF FRAUNHOFER'S LINES

A	7593·8 Å	O	C_2	5172·7 Å	Mg
a	7184·5	H_2O	F	4861·3	Hβ
B	6867·2	O	G	4307·7	Ca
C	6462·8	Hα	*h*	4101·7	Hδ
D_1	5895·9	Na	H	3968·5	Ca
E	5270	Fe	K	3933·7	Ca

continuous light coming from the deeper layers of the sun or star being partly absorbed by incandescent gas. Table II gives the elements to which the various Fraunhofer lines are due. We see that gases are not only capable, as shown above, of emitting light of perfectly definite wave-length when excited to luminosity (**emission spectra**), but also of absorbing light of perfectly definite wave-length (**absorption spectra**).

Thus, as may be shown experimentally (Vol. V), the D line in the sun's spectrum is due to sodium, with whose yellow emission line it exactly coincides in wave-length. A line (D_3) very near the D line is due to helium.† As optical instruments have been perfected, the number of Fraunhofer lines known has risen to many thousands. Some of them are of terrestrial origin (e.g. A, *a*, B), being due to absorption by the atmosphere (many vary with its humidity, the so-called "rain bands").

Band Spectra.—In emission and absorption spectra the lines often appear to be crowded together at certain places, so that in small

* Gases under extremely high pressure also emit a continuous spectrum.

† By this line in the light of solar prominences the element helium was recognized as distinct from sodium in 1868 by the British astronomer NORMAN LOCKYER (1836–1920, director of the Solar Physics Observatory, South Kensington); it was not until twenty-six years later that RAMSAY showed by means of the same line that helium exists on the earth also.

spectroscopes the region appears continuous (cf. the third, fourth. and seventh figures from the top of the coloured plate). These regions where the lines are crowded together are called **bands.** If, however. the spectrum is spread out by using instruments of sufficiently high resolving power, most of these bands are resolved into separate lines (figs. 12, 13, Plate XIV).

The difference between line spectra and band spectra, however. is not entirely an external one but has a physical basis in addition: for, as will be shown in Vol. V, line spectra are emitted by *atoms*, band spectra by *molecules*.

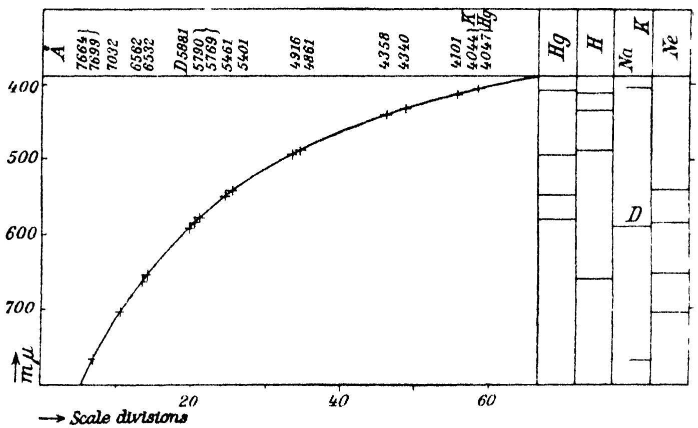


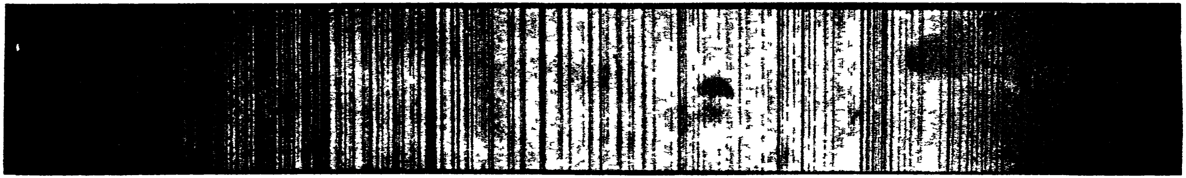
Fig. 14.—An example of the calibration curve of a spectrograph

Calibration of a Spectrum.—The wave-lengths of a large number of lines have been determined very accurately by interference experiments (pp. 15, 208). By means of these known wave-lengths it is very easy to graduate a spectroscope by determining the position of the lines of a known substance and thus associating a definite wave-length with every point on the scale, the relationship between the wave-length and the numbering of the scale being represented graphically (fig. 14) or by mathematical interpolation. For accurate work a spectrum very rich in lines must be used, e.g. that of the arc between iron electrodes.

Other Forms of Spectrometer.—Besides the prismatic spectrometers discussed in this section, interferometers are also in use, especially for more accurate investigations; these will be considered elsewhere (Chap. VIII, §§ 6, 7, and 9, pp. 190, 193, 205).

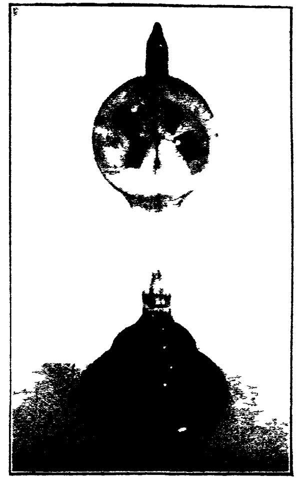
4. Ultra-violet and Infra-red Light.

Ultra-violet Rays.—If the spectrum of the sun or of an arc lamp is allowed to fall on a screen covered with a layer of Sidot's hexagonal

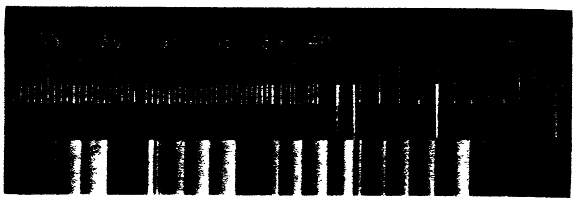


Ch VII Fig 11 Spectrum of Sunlight showing Fraunhofer's lines

Ch VII Fig 13 Band spectrum as obtained with apparatus of high resolving power (Photograph by the Einstein Observatory)



Ch VII Fig 15 — Crookes light mill (p 163)



Ch VII Fig 12 — Band spectrum

zinc blende,* we see that in the violet region of the spectrum the screen shines with a brilliant green light (see *fluorescence*, Vol. V). If we cut off the light from the source, the region of the screen where the blue and violet parts of the spectrum were previously situated continues to shine with its own green light (*phosphorescence*, Vol. V). On closer examination we see that the fluorescence extends beyond the visible violet portion of the spectrum. In regions where no light is visible to the eye, therefore, there is still radiation present which is capable of exciting fluorescence. The radiation beyond the violet is called *ultra-violet light*. This is the light which acts most strongly on the photographic plate, so that it may readily be detected by this means. The light of the iron arc and of the mercury lamp (Vol. III, p. 356) is particularly rich in ultra-violet rays.

The wave-length of this light is determined photographically by means of interference phenomena; it is found to be *less* than that of visible light.

The dark-adapted eye is capable of detecting intense ultra-violet light as a colourless grey. If a direct-vision spectroscope is sighted on diffusely illuminated regions of the sky close to the bright disc of the sun, the portion of the spectrum visible to the eye is considerably extended, to beyond the H line. The Fraunhofer lines K ($\lambda = 393$ $m\mu$) and L ($\lambda = 383$ $m\mu$) can then be observed very well.

Glass absorbs ultra-violet rays beyond about 340 $m\mu$. The Uviol glass made by SCHOTT of Jena is transparent for even shorter waves: to go farther (to about 200 $m\mu$) prisms and lenses of quartz must be used; transparency down to 185 $m\mu$ is obtained in air with fluorite (sometimes also with quartz).

Glasses which *transmit* practically nothing but ultra-violet light (ultra-violet filters) have recently become commercially available. These are much used in practice for the observation of fluorescence phenomena without disturbance from the source of light exciting the fluorescence (quartz analysis lamp).

Ultra-violet rays have a *biological* effect, the region from 320 to 280 $m\mu$ being the most beneficial, the intensity remaining the same, while the shorter wave-lengths have a destructive effect on the tissues. As glass absorbs even the longer wave-lengths, special glass † must be used for windows, &c., if the ultra-violet rays in ordinary daylight are to be utilized. The ultra-violet rays appear to be of additional biological importance in that they evidently transfer their energy, as in the phenomenon of phosphorescence mentioned above, to definite substances which have an important bearing on the life of the organism.

V. SCHUMANN of Vienna showed in 1893 that the absorption of short waves in air makes further advance in the short-wave direction very difficult. He succeeded in detecting ultra-violet rays in the spectrum of hydrogen of wave-length $\lambda = 120$ $m\mu = 1{\cdot}2 \times 10^{-5}$ cm. (Schumann region of the ultra-violet) photographically by means of a vacuum spectrograph, using optical parts of fluorite

* This consists of zinc sulphide containing traces of copper, as well as added chlorides, ignited as far as possible out of contact with air (see Vol. V): it is well known from its use (mixed with radioactive substances) as luminous paint for clocks and watches.

† [The best-known variety is called Vita-glass (introduced in 1925).]

and dry plates containing no gelatine. Using a high vacuum and a grating spectrograph (p. 208) LYMAN was able (1920) to detect and measure helium lines of wave-length $\lambda = 600$ Å. $= 6 . 10^{-6}$ cm., while in 1924 R. E. MILLIKAN and BOWEN, using the highest vacuum meanwhile attainable, detected and measured metal lines from a spark discharge down to $\lambda = 137$ Å. $= 1{\cdot}37 \times 10^{-6}$ cm. for aluminium.

Recent advances have pushed the limit to about 100 Å. For the resolution of the remote ultra-violet, reflection gratings (Chap. VIII, § 10, p. 205) are generally used instead of prisms, as these rays are very strongly absorbed by all substances. Thus the gap between ultra-violet rays and X-rays (Vol. III, p. 347), which extend to about 150 Å. on the long-wave side, has been bridged. For the prismatic spectra of these very short waves see p. 165 and Plate XV; for grating photographs of these and radiations of still shorter wave-length see Plate XVI and Vol. V.

Infra-red Rays.—Into the spectrum of an arc lamp or of the sun we bring a zinc sulphide screen which has previously been made strongly luminous by intense illumination from an iron arc or diffused daylight (not direct sunlight). We then notice that when the screen is brought into the spectrum the phosphorescence in the red region vanishes (so-called *quenching*). This effect, however, can also be detected beyond the red part of the spectrum, so that an invisible radiation must be present on that side also (the so-called *infra-red* rays). As is again established by interference experiments, the wave-lengths of the infra-red rays are *greater* than those of visible light. To measure the intensity of infra-red rays, we use a sensitive thermocouple or bolometer, or the radiometer effect (see below). To detect their presence qualitatively, it is sufficient to use a thermometer with a long narrow blackened bulb.* All parts of the spectrum give a heating effect which is proportional to the energy-density at the point in question, the electromagnetic radiation being transformed into heat on absorption. Thus the intensity of radiation can be quantitatively determined both in the ultra-violet region and in the infra-red region. If we use an incandescent solid as source of light, the energy-density of the infra-red radiation is actually very much greater than that of the visible or ultra-violet radiation (see fig. 20, p. 31, and the article on temperature radiation in Vol. V).

Glass absorbs the longer infra-red rays; further advances (to about 16 μ or 21 μ) can only be made by the use of prisms and lenses of rock-salt and sylvine respectively, while for very long waves quartz is again transparent (for further details see the article on residual rays in Vol. V). Black paper is also transparent to infra-red rays.

The radiation with longest wave-length obtained from a source of light is that from a mercury lamp observed by RUBENS and VON BAYER ($\lambda = 342\ \mu$, 1911), and by NICHOLS and TEAR ($\lambda = 420\ \mu$ (0·4 mm.), 1925). This wave-length actually exceeds that of the shortest electromagnetic radiation which has been produced by electrical methods ($\lambda = 0{\cdot}1$ mm.; see Vol. III, fig. 47, p. 646).

* It was in this way that W. HERSCHEL discovered the infra-red rays in sunlight in 1800.

The Radiometer Effect.—Apart from the bolometer, the radiometer is the most sensitive instrument for measuring infra-red radiation in particular. The radiometer effect was discovered by CROOKES while he was attempting to detect the pressure due to radiation. His well-known "light mill" (fig. 15, Plate XIV) consists of a small vane of thin sheets of mica blackened on one side, arranged crosswise, and placed in a high vacuum. If the "sails" are exposed to radiation, the latter is absorbed by the black sides facing it and these are heated. The molecules of the surrounding gas falling on these heated surfaces on the average receive a greater momentum on collision than if the surfaces were at their normal temperature. Hence the backward momentum which they impart to the surfaces is greater than that of the molecules of gas which impinge on the other (cooler) sides of the mica sheet. The sails will therefore rotate in the direction of the radiation—as they do e.g. if strong daylight falls on them. If the vane is suspended from a torsion head, the deflection produced on irradiation is a measure of the energy of the incident radiation.

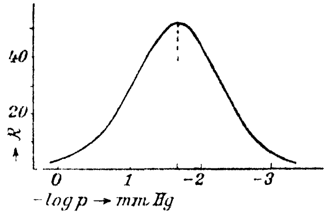


Fig 16.—Variation of the radiometer effect with the pressure inside the vessel

This effect is to be carefully distinguished from radiation pressure, which is not obtained unmixed with other effects unless in an extremely high vacuum, while the radiometer effect is a maximum when the pressure of the surrounding gas has a definite low value (fig. 16). Under certain circumstances, as GERLACH and WESTPHAL have shown, motion may even take place in the direction opposite to that of the incident radiation (see also Vol. III, p. 649).

5. Measurement of Dispersion.

The formation of a spectrum when light passes through a prism is due to the fact that the refractive index varies with the wave-length of the light. The relationship between them is shown for several substances in fig. 17 (the so-called dispersion curve).

By using the equation for δ (p. 62, or for approximate work that for small-angled prisms will suffice) we may calculate the deviations of the various rays for a prism of given refracting angle from the values of μ given in fig. 17 and hence sketch the whole spectrum. This is done in fig. 18 for prisms of four different substances, the refracting angle being the same in each case: water, crown glass, flint glass, and carbon disulphide. The spectra are drawn one above the other in such a way that the C line occupies the same position in all four. We see that the length of the spectrum varies with the material of which the prism is made.

The difference of the refractive indices of a material for the lines H and C is called the **specific dispersion** θ; $\theta = \mu_H - \mu_C$.

The difference of the refractive indices for any other two wave-lengths is called the **partial dispersion** for the region in question. The dispersion for the

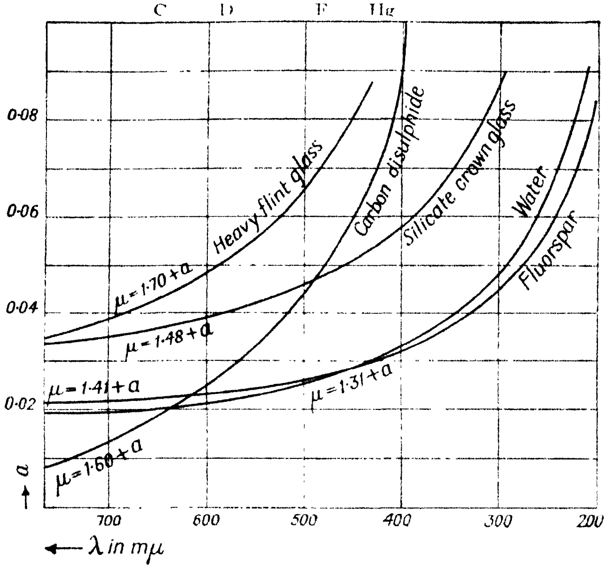


Fig. 17.—Dispersion curves for various substances (the ordinate merely gives the additive term a)

brightest part of the spectrum between the C and F lines is also called the **mean dispersion.** The ratio of this difference $\mu_F - \mu_C$ and the refractive index for the D line diminished by unity is the **dispersive power** of the refracting substance. It is usual to follow ABBE and quote its reciprocal,

$$\nu = \frac{\mu_D - 1}{\mu_F - \mu_C}.$$

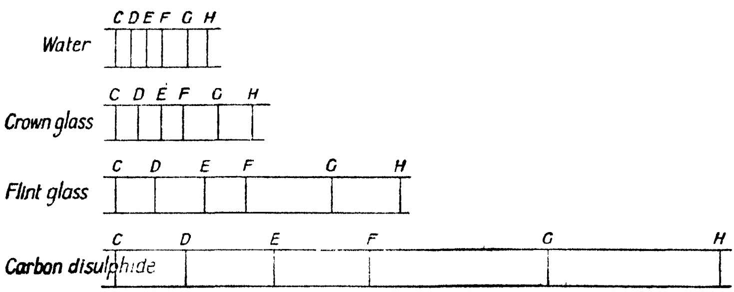


Fig. 18.—Lengths of spectra produced by different substances under otherwise similar conditions

This expression (ABBE's number) is very suitable for characterizing those properties of a substance which have an essential bearing on its use in optical systems.

The total dispersion of a prism of small refracting angle α is

$$\Theta = \theta\alpha.$$

The values of the refractive index given in fig. 17 for light crown glass and heavy flint glass are examples from two types of glass, of which the first has a very small specific dispersion and the second a very large specific dispersion. They merely serve as examples, for nowadays a large number of glasses with extremely varied optical properties are manufactured.

Refractive Index for X-Rays.—For the rays of about 100 Å. and less mentioned on p. 162 the refractive indices of almost all substances are slightly *less* than unity. The difference from unity is extremely small (cf. legend, fig. 20, Plate XV). A consequence of this is that when a beam of such radiation meets a solid body in air the ray is in certain circumstances totally reflected. The ray

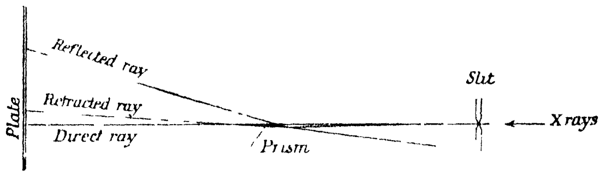


Fig. 19.—Refraction of X-rays by a glass prism according to Larsson, Siegbahn, and Waller (from *Zeitschr. f. techn. Physik*, Vol. 8 (Barth, Leipzig)

must, however, fall on the boundary at almost grazing incidence, as the critical angle differs from 90° by only a few minutes. Owing to this property of the refractive index, these rays are refracted *towards* the refracting edge e.g. of a glass prism. As in this region dispersion still occurs, i.e. the refractive index varies with the wave-length, it is possible to obtain prismatic spectra in this way, as is shown for X-rays of certain wave-lengths in figs. 19 and 20.

6. Achromatic Prisms and Lenses.

Achromatic Prisms.—For a prism of refracting angle a the deviation δ_C for the Fraunhofer C line and the total dispersion Θ are as follows:

$$\delta_C = 0{\cdot}5160a;\ \Theta = 0{\cdot}0089a \text{ for crown glass,}$$
$$\delta_C = 0{\cdot}6144a;\ \Theta = 0{\cdot}0171a \text{ for flint glass.*}$$

It follows that whereas two prisms with the same refracting angle, one of crown glass and one of flint glass, give rise to very nearly the same deviation, the dispersion and hence the length of the spectrum produced is nearly twice as great with flint glass as it is for crown glass. On the other hand, it is possible to make prisms with the same dispersion using either kind of glass: their refracting angles must then be in the inverse ratio of the specific dispersions, i.e. in the ratio 89 : 171 or approximately 1 : 2. Then, however, the deviations produced by the two prisms are approximately in the ratio of 1 : 2 for the C line. Hence if two such prisms are fitted together in such a way that their refracting angles point in opposite directions, the resulting compound prism deviates the rays almost without dispersing them at all. A prism of this kind is called an *achromatic prism.*

* These values are for the types of glass ordinarily used for the purposes described below.

Fig. 21 shows an achromatic prism consisting of a crown-glass prism with a refracting angle of 60° and a flint-glass prism with a refracting angle of 35°. A

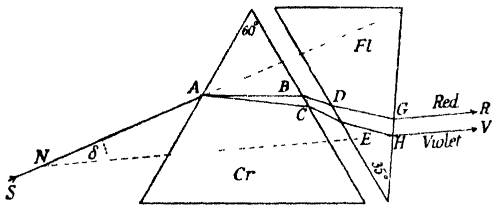


Fig. 21.—Achromatic prism

ray falling on the crown-glass prism from the left is split up into a fan of colours whose outermost boundaries are AB (red) and AC (violet). The dispersion is then counteracted by the flint-glass prism and the two rays GR (red) and HV (violet) travel on parallel to one another in a direction making the angle δ with the original direction. If a whole pencil of rays falls on the prism from S instead of a single ray, all the coloured components emerging parallel to one another are reunited to form white (except for the narrow regions at either extremity; these, however, may be stopped out).

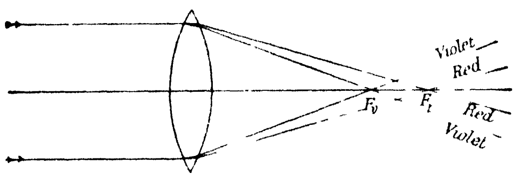


Fig. 22.—If the lens is not achromatic the various colours have different foci

Achromatic Lenses.—We have already explained the action of a lens by regarding it as composed of a number of thin prisms. It follows that it is also possible to combine lenses of crown glass and flint glass in such a way that the combination behaves like a simple lens as regards deviation of the rays, but exhibits no chromatic dispersion. Such combinations are called **achromatic lenses**; they consist of a biconvex lens of crown glass and a concavo-plane lens of flint glass and act like a simple convexo-plane lens, without dispersion. The objectives of telescopes consist entirely of achromatic lenses. If telescopes were made with ordinary lenses, the images of objects would have coloured margins.*

* The possibility of achromatic prisms and lenses occurred to NEWTON, but unfortunately he chose for his experiments two substances of equal dispersive power, and concluded that achromatism is impossible. Achromatic telescopes were first made about 1733 by an English amateur, CHESTER MOOR HALL (1703–1771). His instruments were excellent, but he published no account of them, and the chief credit for the invention must be given to JOHN DOLLOND, a London optician of Huguenot descent who, influenced by the work of NEWTON, EULER and the Swede KLINGENSTIERNA, from 1757 manufactured achromatic telescopes on a large scale.

While in the case of an ordinary lens the violet rays of a parallel pencil incident on the lens are brought to a focus at F_v (fig. 22), which

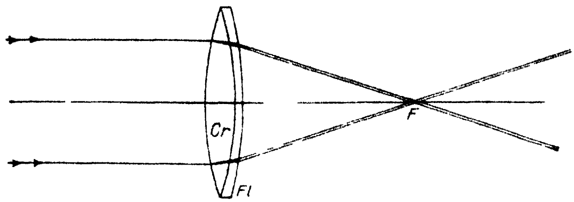


Fig. 23.—Achromatic lenses

lies nearer the lens than F_r, the focus for red rays (**chromatic aberration**), in the case of an achromatic lens formed by the combination of a convergent lens of crown glass and a divergent lens of flint glass, the coloured components emerge parallel to one another and hence

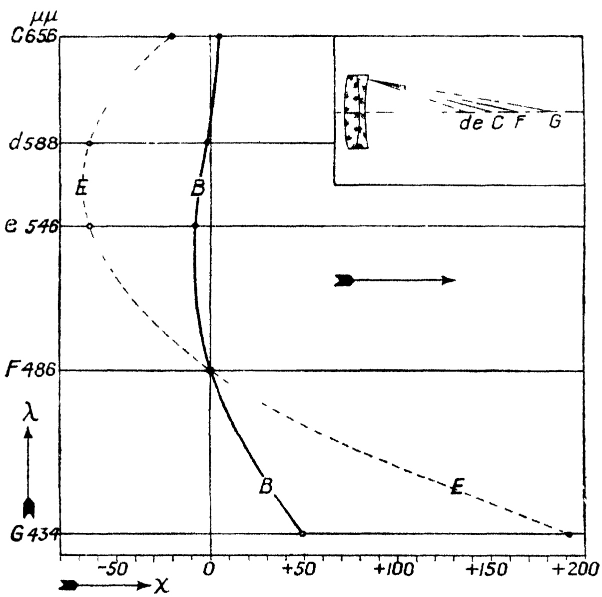


Fig. 24.—Colour curve for two objectives: E is an ordinary objective, B a three-component apochromat. The abscissæ represent hundred-thousandths of the focal length. Above, a diagram of the positions of the foci for five different wave-lengths (Carl Zeiss, Jena.)

are brought to a focus at the same point F (fig. 23). How completely achromatism can be attained is shown in fig. 24 for two astronomical objectives.*

* [In this figure the focal length for the F line is taken as a standard, and the curves exhibit the deviations which occur for other wave-lengths. The corresponding curve for a perfectly achromatic objective would be the vertical straight line through O.]

Secondary Spectrum.—In the combinations of prisms and lenses which we have described, achromatism, strictly speaking, exists only for two arbitrarily chosen points of the spectrum (in the above examples for the C and H lines), the removal of the dispersion being incomplete for the other regions of the spectrum owing to the variations of the partial dispersion; the chromatic dispersion still remaining is called the *secondary spectrum* of the combination. By combining a greater number of lenses of various kinds of glass, the secondary spectrum may also be largely got rid of if the types of glass are suitably chosen. In the microscope apochromats (fig. 28, p. 140) the secondary spectrum has been reduced until it can no longer be detected.

Further, achromatism is only achieved for the rays which traverse the lens parallel to its axis; hence coloured margins are seen even in a telescope (and particularly in the case of ordinary opera-glasses) if one looks obliquely through it.

Chromatic Aberration of the Eye.—Even the human eye is not completely achromatic. If one looks at an upper bar of a window with the head bent far down, the lower edge of each bar will appear red and the upper edge blue. Again,

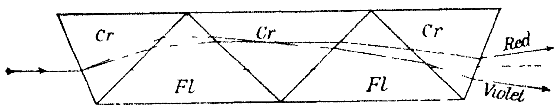


Fig. 25.—Path of the rays in the direct-vision prism

if one pricks a small hole in a piece of paper with a needle and observes a small object, not in the centre of the field of view, through this hole, it will appear to have coloured margins.

Direct-vision Prism.—If two prisms, one of crown glass and the other of flint glass, are combined in such a way that their refracting edges are turned away from one another, and their refracting angles are such that the mean deviations of the two prisms cancel one another, the dispersion of the flint-glass prism is approximately twice that of the crown-glass prism. The central part of the pencil of rays then leaves the combination in approximately the same direction as on entering; the combination forms a **direct-vision prism.***

As a rule three or five prisms are combined as in fig. 25. The various surfaces are cemented together with Canada balsam is order that reflection at the boundary surfaces may be avoided as far as possible. The yellow part of the spectrum in fig. 25 leaves the prism undeviated, the red part being deflected upwards and the violet part downwards. The well-known small **direct-vision spectroscopes** for use in the hand consist of a tube containing a direct-vision prism, an adjustable slit, a collimating lens, and a small telescope.

7. Anomalous Dispersion.

The relationship given by the dispersion curves of fig. 17, p. 164, is the *normal* one for colourless substances in the visible region; the refractive index increases as the wave-length diminishes. Materials exist, however, in which the reverse is the case, i.e. in which the refractive index diminishes as the wave-length does so, or, in general,

* Direct-vision prisms were first made by the Italian physicist **Amici (1786–1863)**; hence they are often called *Amici prisms*.

in which the dispersion curve exhibits maxima and minima. This is actually the case for *all* substances if we consider a wider range of the spectrum than the visible region. The phenomenon is known as **anomalous dispersion.**

Experimental Facts.—Anomalous dispersion was first observed by LE ROUX in 1861 with iodine vapour; the phenomenon was thoroughly investigated by CHRISTIANSEN * in 1870. He determined the refractive index of an 18·8 per cent solution of fuchsine in alcohol by means of a liquid prism of this substance with a refracting angle of 1° 14′. The results of his experiments for a number of the Fraunhofer lines are as follows:

B	C	D	E	F	G	H
1·450	1·502	1·561	—	1·312	1·285	1·312.

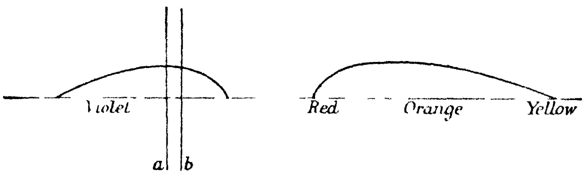


Fig. 26 —Anomalous spectrum obtained by Christiansen by means of a prism containing fuchsine solution

Fig. 26 is a graphical representation of the spectrum obtained by CHRISTIANSEN with this prism. The abscissæ denote the deviations and the ordinates are proportional to the intensities of the colours; the small interval *ab* gives the length of the spectrum produced by the alcohol used for the solution, the refracting angle being the same. By comparing this interval with the length of the whole figure, we gain an idea of the extraordinary magnitude of the dispersion of the solution of fuchsine in alcohol ($\delta = 0{\cdot}276$) as compared with that of alcohol ($\delta = 0{\cdot}013$).

KUNDT found that all substances with surface colour, especially those with a metallic lustre, exhibit marked anomalous dispersion. In solution these substances absorb a fairly sharply-defined region of the spectrum; in the neighbourhood of the absorption region the deviation from the normal behaviour is particularly well marked. This behaviour is also exhibited by fuchsine, whose alcoholic solution absorbs the green almost completely, even in very thin layers (cf. fig. 26). In the solid state fuchsine has a pronounced greenish-gold metallic lustre, which is due to the fact that the greenish-yellow rays which are strongly absorbed by it are also strongly reflected.

In his investigations KUNDT used the *method of crossed prisms*. A beam of white light (fig. 27) falls on the prism I, which has its refracting edge K_1K_1 horizontal; it is thereby spread out into a fan of colours in a vertical plane, which would produce the spectrum AH on the vertical screen S if the prism II with its refracting edge K_2K_2 vertical were not interposed in the path of the rays. Each of the rays giving rise to the fan of colours is refracted by the second prism. Then if the second prism is of the same material as the first, the violet part of the fan of colours, which would otherwise be found at H, is most deviated; while the red part, which would be found at A, is least deviated, so that the oblique rectilinear spectrum A′H′ is produced on the screen (see also fig. 28). If, however, the two prisms are of different materials, the combined effect of the two prisms

* CHRISTIAN CHRISTIANSEN (1843–1917), Professor of Physics at Copenhagen from 1881 to 1912.

does not give rise to a straight spectrum unless the refractive indices of the two substances are proportional to one another for all parts of the spectrum. If this

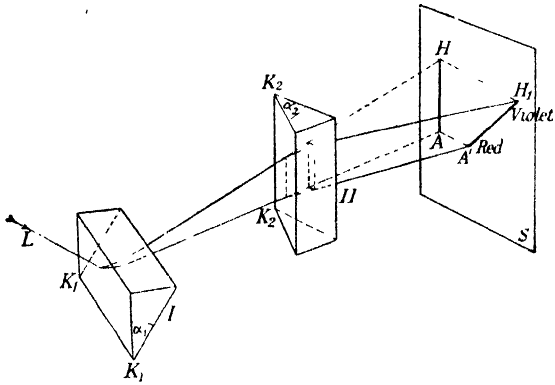
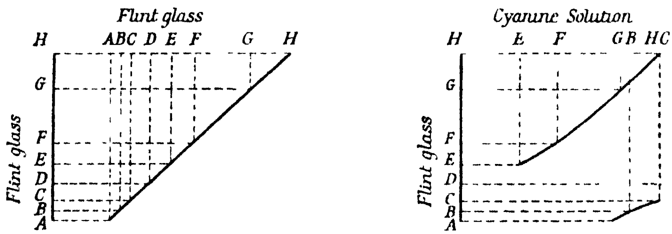


Fig. 27 —Kundt's method of crossed prisms

is not the case the spectrum is curved. Thus if two prisms, one of crown glass and the other of flint glass, were crossed, they would give rise to a curved spectrum in fig. 28.

Fig. 29 shows the spectrum produced when a flint-glass prism and a prism



Figs. 28, 29 —Spectra obtained with crossed prisms

containing concentrated cyanine solution are crossed. The region round the D line is completely absorbed by cyanine; the neighbouring region on the one side, corresponding to the C line of the spectrum, exhibits stronger anomalous dispersion than the neighbouring region of the E line on the other side, although in the normal spectrum this lies much closer to the violet end than the C line does.

Kundt also succeeded in producing extremely thin wedge-shaped sheets of metal, which enabled him to determine the refractive indices of metals for various regions of the spectrum. He thereby established the fact that the refractive indices of metals, e.g. of silver, gold, and copper, for sodium light are *less* than unity (e.g. 0·27 for Ag, 0·58 for Au, 0·65 for Cu in yellow light); further, that a simple relationship exists between their refractive index (especially for long wave-lengths) and their electrical conductivity. These experiments were subsequently repeated by Du Bois and Rubens * and also by Hagen and Rubens and extended in a variety of ways.

* Heinrich Rubens (born in 1865 at Wiesbaden, died in 1922) was made Professor of Physics in Berlin in 1906; he was previously at the Technische Hochschule, Charlottenburg.

In 1880 KUNDT went on to observe the phenomenon represented in fig. 30, which occurs when a bunsen flame full of sodium vapour is placed in the path of the rays forming the spectrum of white light. The dark D line produced by absorption of the light by the incandescent sodium vapour exhibits distortion as a result of anomalous dispersion. The sodium flame acts as a prism with its refracting edge vertical, so that we really have the crossed prism arrangement again.

In 1898 HENRI BECQUEREL * repeated the experiments on the anomalous dispersion of incandescent vapours, using a spectrometer with a dispersion sufficient to show the D line as two lines (p. 158). As a result he discovered the phenomenon illustrated in fig. 31. In the neighbourhood of each of the two absorption lines the refractive index of sodium vapour is extremely large for the rays on the red side, but less than unity for the rays on the violet side.

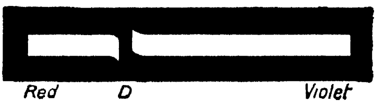


Fig. 30

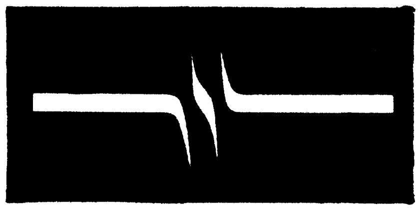


Fig. 31

Anomalous dispersion of sodium vapour

Explanation of Anomalous Dispersion and of Normal Dispersion.—The cause of dispersion was first recognized by SELLMEIER in 1871. He assumed that the molecules have proper vibrations of definite frequencies. The resulting resonance phenomena give rise (as STOKES had also assumed earlier) to absorption, for the energy transferred to the resonators and used up in their vibrations is obtained from the incident radiation; further, the velocity of propagation of light is also affected. This varies according to the difference between the frequency of the light and the frequency of the excited proper vibration of the molecule, so that dispersion occurs.

The Propagation of an Electromagnetic Wave through a Material Medium.—It is essential that we should consider this process, in order that we may understand refraction phenomena. It is thought of as follows: each material particle (atom or molecule) carries an electric charge (Vol. III, p. 641). It is therefore capable of executing electromagnetic vibrations (of being a resonator); let its proper frequency be ν_0. (Usually there are many proper frequencies; see below.) On p. 224 of Vol. II we have already shown for wave motion generally that if the system has a proper vibration the velocity of propagation will differ for different wave-lengths, for the incident

* HENRI BECQUEREL (1852–1908) was latterly professor at the École Polytechnique in Paris; his father ALEXANDRE EDMOND BECQUEREL (died 1891) and his grandfather ANTOINE CÉSAR BECQUEREL (died 1876) were also well-known physicists.

wave of frequency ν excites the resonators to execute forced vibrations. Each resonator becomes the starting-point for a spherical wave, whose amplitude in the case of undamped resonators is smaller than that of the incident wave in the ratio of $1 : (\nu_0^2 - \nu^2)$; there is also a difference of phase between the two, and the amplitude is affected by damping, if such exists (see the account of resonance phenomena on p. 198 of Vol. II, especially figs. 17 and 19). The exciting wave, however, is itself due (according to Huygens' principle) to interference of the spherical waves which start from the individual resonators in the neighbourhood surrounding the resonator in question. The result of this complicated interference phenomenon is then what we call the ray traversing the medium. The resultant wave therefore depends in a very complicated way on the individual elementary waves which are set up. At each particle the phase of the wave is somewhat altered, so that the velocity of the wave (more accurately, the phase of the wave) is altered (accelerated or retarded) as compared with the velocity in free space. In direct measurements of velocity, what we observe is merely the *group velocity* (Vol. II, pp. 222, 226; see also the present volume, Chap IX, § 2, p. 219). The *wave velocity* (*phase velocity*) is obtained by dividing the velocity of light in free space by the refractive index. From this we must also distinguish the so-called *signal velocity*, i.e. the velocity with which the front of the wave motion advances. As consideration of the phenomenon described above shows, it is independent of the material filling the space; the latter affects the phase only, as a result of the sympathetic vibrations of the electric charges and resonance. In material media the signal velocity is still equal to c, the velocity of light in free space.

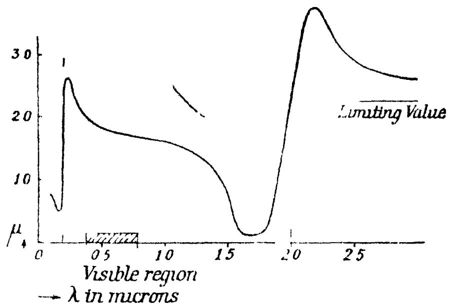


Fig. 32.—Complete dispersion curve (after Jentzsch). Ideal case of a substance with absorption bands at 200 $m\mu$ and 2μ

A number of special cases of the propagation of light will be discussed in rather more detail in Vol. V.

We see from the above discussion that the phase velocity, and hence also the refractive index, depends in each case on the proper frequency of the resonators and their degree of damping. For substances with normal dispersion in the visible region it is frequently sufficient, in order to represent the principal features of the dispersion, to assume only two frequencies, one in the ultra-violet and one in the infra-red. Fig. 32 exhibits a curve obtained theoretically, drawn on the basis of the more detailed formulæ which were given subsequently by HELMHOLTZ,

KETTELER, DRUDE, LORENTZ, and PLANCK. As we see, the dispersion is anomalous on the short-wave side of a proper vibration. We see this in figs. 30 and 31 also, since the D line corresponds, as its absorption in sodium vapour shows, to a resonance frequency of the sodium atom.

On the long-wave side the dispersion is normal; it decreases continuously after reaching a maximum, whose value and position depend on the degree of damping. As we see, the region of "normal" dispersion is only a small section of the more general dispersion curve. Normal dispersion in the visible region is to be ascribed to the effects of absorption in the ultra-violet and usually in the infra-red as well. The ultra-violet vibrations are caused by electrons, the infra-red by ions, as is also revealed by the phenomena of *residual rays* (Vol. V). In accordance with the behaviour of the dispersion curve, the refractive index rises as the wave-length diminishes (i.e. in the ultra-violet, as the resonances due to the electrons often lie very far in the ultra-violet) for almost all substances which are transparent in the visible region. For very short waves (e.g. X-rays) the refractive index loses much of its importance, as the wave-length is then of the order of the dimensions of a molecule (10^{-7} cm.). *Scattering*, such as we observe even in the visible region with relatively large particles (smoke, milk), then plays the principal part (p. 232). The refractive index for X-rays, as we have seen, has been found to be slightly less than unity (p. 165).

It is an interesting experiment to construct an artificial medium with resonators of definite periods and to determine its refractive index in the neighbourhood of the resonance wave-length. An experiment of this kind was made by C. SCHÄFER with electric resonators of wave-length 82 cm., and he was able, using correspondingly long electric waves, to establish the existence of normal and anomalous dispersion on the two sides of the proper frequency.

Molecular Refraction.—The alternating electromagnetic field of the incident light gives rise to polarization of the molecules (Vol. III, p. 99), the charges on the molecules being slightly displaced relative to one another. If we substitute MAXWELL's relationship * $D = \mu^2$ (Vol. III, p. 644) in the equation for the molecular polarization which we obtained on p. 102 of Vol. III, we have the so-called molecular refraction

$$R_M = \frac{\mu^2 - 1}{\mu^2 + 2}\,\frac{M}{\rho} = \frac{N}{3\Delta_0}\,\alpha.$$

We thus have a means of obtaining data on the electrical properties of the medium from optical experiments (see ionic refraction, Vol. V). Even permanent dipoles are able to follow the variations of the alternating field, provided these are not too rapid. To this are due the dispersion phenomena in such substances (Vol. III, p. 104) in the region of short electrical waves (DEBYE).

* Here D and μ are to be taken for the same wave-length, which must be at a sufficient distance from the position of resonance, as MAXWELL's relationship takes no account of resonance, i.e. applies only to the asymptotic portion on the right-hand side of fig. 32.

CHAPTER VIII

The Interference of Light Waves

1. Fresnel's Interference Experiment.

Coherence and Capacity for Mutual Interference.—During the whole of our discussion of optics up to this point, we have had to take account of modifications of the phenomena and of actual interference phenomena due to the undulatory nature of light. Now that we have some acquaintance with optical relationships and the way in which optical instruments work, we shall proceed to investigate the conditions under which light interference occurs in greater detail.

We must again emphasize that the presence of *coherent* wave-trains differing in phase is essential if lasting interference phenomena are to occur. Waves of any energy sent out in any direction by a large emitter of electromagnetic waves are coherent for any length of time (in the case of continuous excitation by a valve transmitter, but not with a quenched-spark transmitter), just like the waves emitted by a uniformly and continuously excited tuning-fork. On the other hand, the electromagnetic waves emitted by a source of light are not coherent for *any* time and in *any* direction. As we have already explained on p. 2, "ordinary" light consists of a sum of mutually incoherent wave-trains, each of which is coherent in itself. As an acoustical model we might adduce a very large collection of tuning-forks, each of which is excited at entirely irregular intervals and gradually comes to rest. The sum of all these separate wave-trains would then be the "ordinary" radiation from the collection of tuning-forks.

That is, only light which starts from one and the same point of the source (probably from a single atom or molecule) can be coherent. Among the immense number of wave-trains emitted by a strong source of light, a coincidence of phase at different points may occur accidentally for a short time. We shall not go further into this phenomenon (which is known as **consonance**), nor shall we discuss the possibility of relationships existing permanently between the phases at different points of the source of light. For some remarks on this point see Vol. V.

Fresnel's Mirror Experiment.—One of the historically most impor-

tant attempts to produce interference phenomena systematically by the combination of coherent wave-trains is Fresnel's mirror experiment.

Two plane mirrors of black glass meet one another along a straight edge in such a way that neither mirror projects beyond this intersection. The two mirrors are inclined to one another at a very small angle (only a few minutes). At a distance of a few centimetres from the common edge a narrow slit is set up vertically, close to the surfaces of the mirrors, in such a way that it is exactly parallel to the common edge. The slit is illuminated from behind by a bright source of light so that the rays of light passing through the slit fall on both mirrors and are reflected by them (fig. 1).

The pencils of rays starting from the slit L and reflected by the two mirrors S_1 and S_2 behave as if they were due to the two virtual images L_1 and L_2. As contrasted with two arbitrary sources of light, they have the important additional property of emitting coherent light.

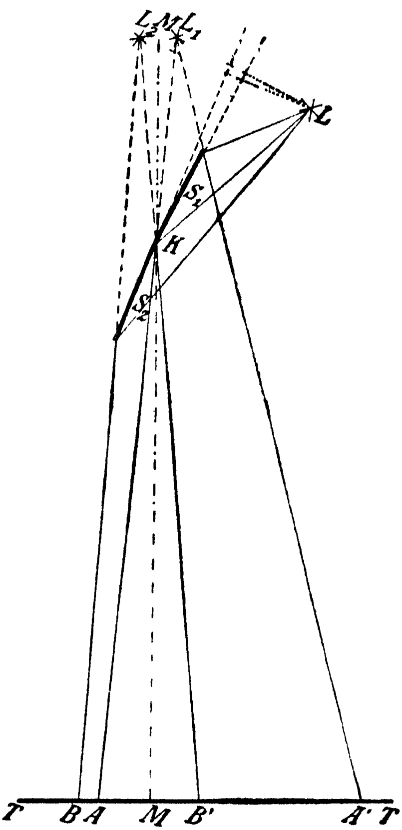


Fig. 1 —Fresnel's mirror experiment

The two pencils of rays cross one another in the region common to them both in the same way as the two systems of waves do in the experiment with water waves (Vol. II, p. 214). In fig. 2 the intersections of the waves are shown as if the wave systems came directly from L_1 and L_2. The heavy circles denote the wave crests, the light circles the wave troughs. These give rise to a family of hyperbolas along which wave crests always coincide with wave crests and also wave troughs with wave troughs. Between these lie hyperbolas along which a wave crest of one system always coincides with a wave trough of the other system. Hence if a white screen is brought into the region where the waves overlap, alternate bright and dark bands (interference bands) will be seen on it, and in particular there must be a bright band exactly in the centre of the interference region, along the line which bisects L_1L_2 at right angles.

On the white screen there is a bright spot of light due to simultaneous illumination by both sources (fig. 3), which is intersected near the centre by bright

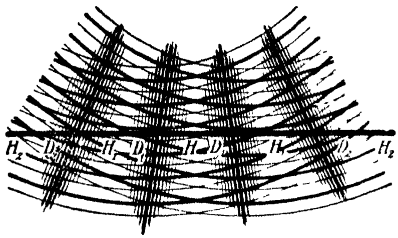


Fig. 2.—Explanation of Fresnel's mirror experiment

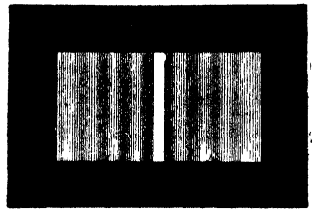


Fig. 3.—System of interference bands in Fresnel's mirror experiment

and dark vertical bands. The bright bands correspond to those points at which the two pencils of light are in the same phase, while at the dark bands the two pencils are in opposite phases. The bright and dark bands seen on the screen are the lines along which the screen intersects the family of confocal hyperbolas; in the case of water waves the latter may also be observed directly.

Fresnel's mirror experiment may be used to determine the wave-length of the light under investigation.

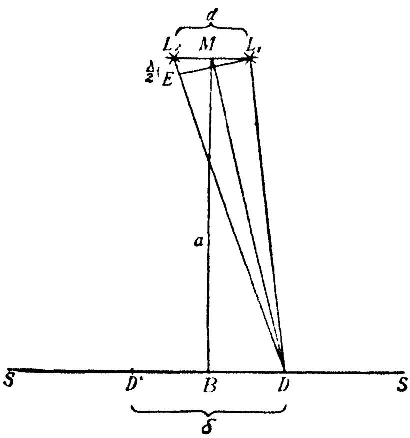


Fig. 4.—To illustrate the theory of Fresnel's mirror experiment

Let L_1 and L_2 in fig. 4 be the virtual images of the slit produced by the two mirrors; L_1 and L_2 are to be regarded as the starting-points of the two systems of waves. Let SS be the cross-section of the screen on which the Fresnel bands are received; it is set up at a great distance from L_1L_2 (in order that the interference hyperbolas may be regarded as straight lines). MB is the perpendicular bisector of L_1L_2 and meets the screen at B; then B is a point in the bright central band, for $L_1B = L_2B$. The point B is a point of the same phase in both systems. The point D lies on a dark band if the difference of its distances from L_1 and L_2 is half a wave-length. D′ denotes the corresponding dark band on the other side of B.

We put $L_1L_2 = d$, $DD' = \delta$, and $MB = a$. As on p. 16, we have

$$\lambda = \frac{\delta d}{a},$$

where a and δ can be measured directly. To determine d we interpose a convex lens in the path of the rays as shown in fig. 5 and thus produce real images $\mathcal{L}_1$, $\mathcal{L}_2$ of the virtual point sources L_1, L_2 on a screen provided with a scale. We can then find m, the distance between the slit and the lens or (to a sufficient degree

of accuracy) the distance between the apparent sources of light and the lens, n, the distance between the real images and the lens, and b, the distance between the real images. By geometry it then follows that $d = mb/n$, so that all the quantities required to determine λ are measurable.

Example.—$a = 2000$ mm., $\delta = 2$ mm., $m = 220$ mm., $n = 2200$ mm., $b = 5{\cdot}9$ mm. (the focal length of the convex lens is 200 mm.). Hence $d = mb/n = 0{\cdot}59$ mm., $\lambda = 0{\cdot}59 \times 2/2000 = 0{\cdot}00059$ mm. $= 590\ m\mu$.

Fresnel also succeeded in producing two coherent sources of light by means of two prisms with small refracting angles placed base to base (biprism); the joint effect of these sources again gives rise to Fresnel's interference bands.

Coloured Interference Bands.—If we use monochromatic light, e.g. the filtered green light from a mercury arc lamp, the system of bands is green and black; if, however, we use white light, the bands are

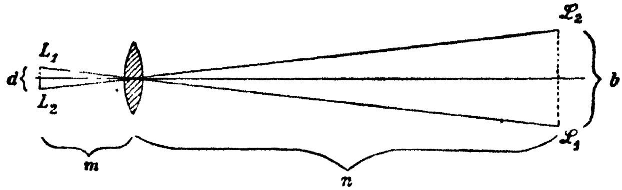


Fig. 5.—The measurement of the distance $L_1 L_2$ (d)

coloured. This is due to the fact that, as we have seen, the breadth of the bands is different for every wave-length, so that the various monochromatic phenomena overlap (see below). If we let the system of interference bands fall on the slit of a spectroscope so that the bands are at right angles to the slit, we obtain the figure shown on the coloured plate (facing p. 158), a spectrum crossed by dark curved bands; from this it at once follows that the wave-length of light is greater the nearer it lies to the red end of the spectrum (pp. 16, 208).

If the interference bands are very close together, we only see the coloured margins of the bands. If, on the contrary, the bands are far apart, each bright band takes the form of a complete spectrum.

2. The Colours of Thin Films.

The well-known iridescent colours of soap-bubbles, the variegated colours of thin layers of oil spread out on the surface of water, and the colours which appear at cracks in colourless glass are all due to the interference of light.

Experiment 1.—If a plane rectangular framework of wire about 5 cm. by 5 cm. is dipped into soap solution and then lifted out again, parallel horizontal coloured bands are seen on the soap film stretched across the frame, provided the latter is held vertically in such a way that two of its edges are horizontal. The bands are found to move gradually downwards and farther apart, while the upper

portion of the film becomes thinner and thinner and finally tears apart. Just before the film tears, its upper portion becomes colourless and transparent; by transmitted light it appears bright, by reflected light dark. If the experiment is repeated with the film illuminated by yellow sodium light, the bands are uniformly bright or dark and are exhibited in large numbers and with great distinctness.

Experiment 2.—Two perfectly flat pieces of plate glass 20 cm. by 5 cm. are laid one on top of the other in such a way that they are in complete contact along one of the short edges, while at the opposite edge they are separated by a thin piece of paper or a strip of aluminium foil. Thus a thin wedge-shaped

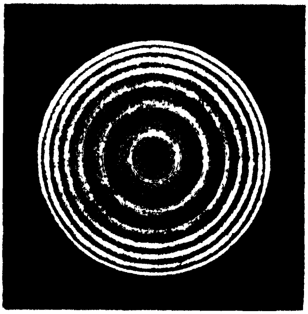


Fig. 6.—Newton's rings

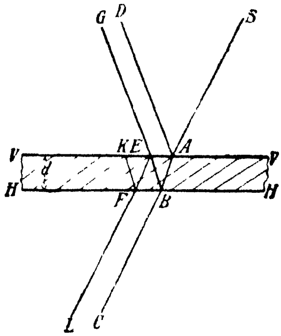


Fig. 7.—To illustrate the theory of colours in thin films

air-space is left between the plates; when it is illuminated by white light, coloured bands are seen, especially at the end where the plates are closest together. If sodium light is used, the whole air-space is crossed by bright and dark parallel lines from end to end.

Experiment 3.—If a slightly convex lens, e.g. a spectacle glass of focal length 4 m., is laid on a flat piece of plate glass, the point of contact of the lens and the plate is surrounded by a system of many-coloured concentric circular rings, which on illumination by sodium light are transformed into a system consisting of a large number of bright and dark rings (fig. 6).

The common feature in all three experiments is the occurrence of coloured or bright and dark interference bands when light falls on a thin film of substance with a refractive index differing from that of the surroundings; hence the phenomena are referred to as the **colours of thin films.** The third phenomenon was first observed scientifically and investigated by NEWTON (1676); hence this phenomenon is referred to as **Newton's rings.***

These phenomena may be explained as follows.

* NEWTON, who inclined to the corpuscular emission theory of light, attempted to explain the occurrence of colours in thin films by a special property of the particles composing the ray of light. According to NEWTON, the minute corpuscles falling on a refracting surface receive a disposition which recurs at regular intervals as the ray passes on; this enables the particles to penetrate readily into the next refracting surface at each return, but to be readily reflected *between* each return. These periodically recurring dispositions NEWTON called fits of easy transmission and fits of easy reflec-

Let a parallel pencil of light fall on a thin layer or film bounded by parallel planes. Then on both the front and back surfaces of the film the pencil of rays is split up into a reflected part and a transmitted part. The two parts of the pencil which are reflected from the front and back surfaces can then interfere with one another, and so can the parts which are transmitted.

In fig. 7 VVHH denotes a section of the thin film, e.g. of a soap bubble, the dimensions being greatly exaggerated. Let there be air above and below the film. A light ray SA forming part of a parallel pencil falls almost vertically on the front surface of the film (in the figure the ray is drawn obliquely in order that its various parts may be separately visible). At A part of the ray is reflected along AD by the front surface of the film, while part of the ray enters the film, travelling along AB. At B partial reflection again takes place, one portion of the light travelling along BE, and another portion emerging into the air along BC. The last reflected portion is again partially reflected into the film along EF at the front surface, another portion entering the air and travelling in the direction EG. The portion EF is then partly reflected again at F on the back surface, the light travelling along FK, and another portion FL entering the air. The last reflected portion FK is now again reflected at the front surface. Thus the whole ray is split up into two systems of rays, one system going back into the original medium and the other entering the medium behind the film. If the light is incident almost at right angles to the film, AD, EG, &c., and also BC, FL, &c., fall in the same line. We now consider only the rays AD and EG in front of the film and the rays BC and FL behind the film, as these are far more intense than the others and, moreover, the relationships which hold for these rays are also valid for the remaining rays.

If the thickness of the film is d, the optical difference of path (p. 103) between the rays BC and FL is BE + EF. If this difference of path amounts to half a wave-length or an odd multiple thereof, the two rays BC and FL, if equally intense, would completely cancel each other, i.e. by transmitted light the film here would appear dark. As, however, the intensities are not equal, the intensity of the light is considerably enfeebled but not extinguished. If the rays are incident on the surface at right angles, the difference of path is $2\mu d$.

It follows that at all places where the optical thickness of the film (μd) amounts to a quarter of a wave-length or an odd multiple thereof, the film must appear dark by transmitted light. The two rays AD, EG which return into the front medium also have a difference of path of $2\mu d$ when the light is incident on the film at right angles. Hence if this difference is an odd multiple of half a wave-length, i.e. if the optical thickness of the film is an odd multiple of a quarter wave-length, the film must also appear dark by reflected light. It would follow that the film should appear dark by reflected light at all the points where it appears dark by transmitted light. This, how-

tion. From the coloured rings he calculated that for yellow light the interval between the fits was 1/890,000 of an inch. According to NEWTON, the dark rings by reflected light are due to the fact that the thickness of the layer is equal to a multiple of the interval between the fits. The light which passes through the first surface of the layer during the fit of easy transmission is then in the fit of easy transmission again when it meets the second surface of the layer, and hence is not reflected. Nowadays we are gradually returning to ideas similar to those of NEWTON (Vol. V).

ever, is contrary to observation. The film is bright by reflected light at those points where it is dark by transmitted light. Hence some other circumstance must play an important part in the phenomenon.

We have already shown in Vol. III (p. 636) that fairly long electromagnetic waves are subject to a change of phase on reflection at a metallic wall (an optically denser medium). The present phenomenon shows that this is also true for the electromagnetic waves which form light (cf. Vol. III, p. 642).

The ray SA which enters the film does not thereby suffer a change of phase, nor does the ray BC which emerges at B; the ray BC therefore leaves the film without change of phase. The ray BE reflected at B is subject to no change of phase at the optically less dense medium; at E it is reflected at the optically less dense medium and again suffers no reversal of phase; finally it travels along FL in the space behind the film without a reversal of phase. The relationship deduced above accordingly holds unchanged for the light passing through the film, that is, the film appears dark by transmitted light at those points at which its thickness is an odd multiple of a quarter wave-length.

The light which returns into the first medium behaves differently. The ray SA is partly reflected back at the optically denser medium; the reflected part AD suffers a change of phase, i.e. a change of path of half a wave-length. The ray EG which arises from reflection at B, i.e. at the optically less dense medium, suffers no change of phase. If the film were infinitely thin at this point, complete darkness would arise by the interference of AD with its reversed phase and EG; the thickness $d = 0$ accordingly corresponds to a dark region of the film in reflected light. If the difference of path $AB + BE - 2\mu d$ is such that the phase of AD is altered by a whole wave-length or a multiple thereof, the phenomenon is not altered in the slightest. Hence it follows that at the points where the optical thickness of the film (μd) is a multiple of half a wave-length or, what is the same thing, an even multiple of a quarter wave-length, the reflected rays are extinguished, that is, at these points the film appears dark by reflected light.

The observed fact that the film appears dark by reflected light at points where it appears bright by transmitted light, and conversely, accordingly demonstrates the correctness of our assumption that light waves have their phase reversed on reflection at an "optically denser" medium.

The **Newton's rings** formed between a convex lens and a flat piece of plate glass are only seen very indistinctly in ordinary daylight, the central portion alone being visible. The colours of the rings cannot be accurately measured.* If, however, Newton's rings are observed by monochromatic light (p. 4), the whole space between the lens and the plate is traversed by black rings up to the edges of the lens.

A convenient way of producing a monochromatic light is to arrange a brass tube so that it may be slid up and down the tube of a bunsen burner; at its upper end the brass tube ends in two flat pieces of metal which are bent away from one another. On the top edges of these flat strips is laid a piece of asbestos board 1 cm. across, which is saturated with a solution of common salt. The flame then spreads past the edges of the asbestos, giving a broad flame which emits practically nothing but monochromatic yellow light of wave-length $\lambda = 589\ m\mu$.

* NEWTON carefully measured the diameters of the rings (1676). The results were accordingly available to THOMAS YOUNG, who was the first to solve the problem of determining the wave-length of light quantitatively (1802).

To produce Newton's rings we use a piece of plate glass 45 mm. by 45 mm., on which we place a plano-convex spectacle glass of power 0·25 dioptres. If these are illuminated by sodium light, a large number of distinct rings are clearly visible even to the naked eye. Better still, we may use the small auxiliary apparatus shown in fig. 8, in which the light is reflected vertically on to the lens and plate by means of a glass plate; the rings are observed through a magnifying-glass and may be measured by means of a scale laid on the lower lens.

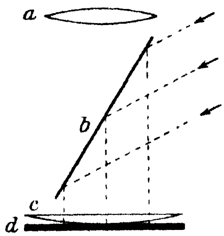


Fig. 8.—Diagram of an apparatus for measuring Newton's rings: *a*, magnifying glass; *b*, glass plate acting as reflector; *c*, lens; *d*, piece of plate glass.

To the distance between two dark rings corresponds a difference in thickness of the layer of air lying between the lower plate and the lens equal to half a wave-length. That is, the difference in diameter of the mth and nth rings corresponds to a difference in thickness in the layer of air of

$$\Delta d = (n - m)\frac{\lambda}{2},$$

since the refractive index may be set equal to unity.

The radius of a ring (a), the thickness of the layer of air at that point (d), and the radius of curvature of the lens (r) are connected by the equation

$$d = \frac{a^2}{2r}.$$

Proof: let FFGG in fig. 9 represent the cross-section of a thick piece of plate glass, and let the dotted segment bounded by HH represent the principal section of a glass lens whose centre of curvature is M. Under the conditions of experiment the lines drawn from the centre of the sphere to the circumferences of the rings include such small angles that arc, sine, and tangent are interchangeable.

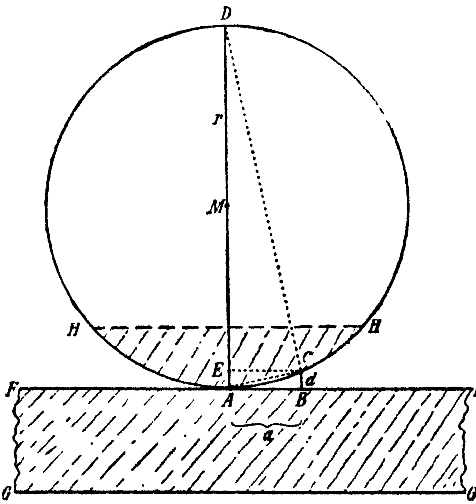


Fig. 9.—To illustrate the theory of Newton's rings

Newton's rings are produced by the interference of rays which are reflected in the same direction, some by the spherical surface of the lens next the air and some by the surface FF. Reflection of a ray falling on the surface HH at this surface contributes nothing to the production of Newton's rings, nor does reflection at GG, the lower surface of the glass plate, as both are too far removed from the thin layer of air. Hence we may neglect the changes of the light ray at the surfaces HH and GG as of no consequence.

Let A be the point of contact of the spherical surface of the lens with the

plate, B a point on a ring of radius a. Let BC, the thickness of the layer of air at B, be d. As construction lines we draw the diameter AD, and also CE perpendicular to AD.

By geometry we have $AE/AC = AC/AD$. Owing to the smallness of the angle ADC, we may replace AC by a. Hence we have $d/a = a/2r$, i.e. the equation $d = a^2/2r$, the truth of which was to be proved.

The radius of the ring (a) is immediately accessible to measurement, while r, the radius of curvature of the lens, may be determined with the help of the spherometer (Vol. I, p. 13) or as on p. 71 of the present volume. The difference in thickness of the layer of air at the mth and nth rings is accordingly given by

$$\Delta d = \frac{a_n^2 - a_m^2}{2r} = \frac{(a_n - a_m)(a_n + a_m)}{2r} = (n - m)\frac{\lambda}{2}.$$

Newton's Rings by Transmitted Light.—Fig. 10 is a diagrammatic representation of an apparatus by means of which the coloured Newton's rings produced in white light both by reflection and by transmission may be projected simultaneously on a magnified scale. N is the combination producing the rings, which consists of a flat plate and a convex lens of slight curvature. This is obliquely illuminated with approximately parallel light by means of a strong source A (an arc lamp) and the condensing lens C. The rays reflected from the lens and plate are recombined by L_1 and the transmitted rays by L_2 in such a way that magnified real images of the various points of the lens and plate are produced on the two screens S_1 and S_2; magnified images of the coloured Newton's rings are thereby produced at the same time. They have a dark centre at M_1 on the screen S_1 and a bright centre at M_2 on the screen S_2. These are surrounded by a number of rings, which are strongly coloured, especially the inner ones. On the screen M_1 the inner edge of the first ring is blue and the outer edge red, while on the screen M_2 the succession of colours is reversed.

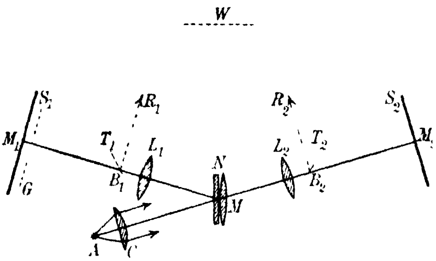


Fig. 10.—Apparatus for observing Newton's rings by transmitted light and by reflected light

If at B_1 and B_2 behind the lenses L_1 and L_2 we set up plane mirrors whose front surfaces are alone reflecting, the rays may be reflected in the directions R_1 and R_2 so that the two systems of rings may be received simultaneously on a screen W where the rays R_1 and R_2 meet (in the figure the screen is not in the right place; it should be farther away from the rest of the apparatus). The superposition of the two systems of rings then produces a colourless patch of light. The colours of the two systems of rings are therefore complementary.

Removing the mirrors at B_1 and B_2 we insert a glass plate, half of which is red and half blue, in front of the screen S_1. We then obtain the configuration shown in fig. 11: the number of rings becomes considerably greater than when no coloured plate was there, and in the red half the rings are farther apart than in the blue half. In fig. 11 the dark rings are indicated by thick semicircles, the bright rings by thinner semicircles. The ratio of the wave-lengths of red light and blue light may be deduced from the ratio of the diameters of corresponding rings. The last experiment at the same time explains why the coloured Newton's rings produced when white light is used and also all other interference phenomena are, as we have said before, distinct only in the neighbourhood

of the centre and grow pale at a short distance from it, finally vanishing completely. We have merely to note that to each separate colour there corresponds a perfectly definite system of rings, the radii of the rings varying from colour to colour. In the neighbourhood of the centre, i.e. where the layer of air is thinnest, the systems of rings corresponding to the various colours are still separate. At a greater distance from the centre, i.e. at places where the layer of air is thick, the systems of rings corresponding to the different colours overlap, giving rise to white. The same is true for all interference phenomena with white light.

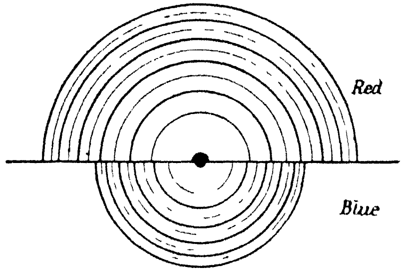


Fig. 11.—Newton's rings with red light and with blue light

Fizeau's Phenomenon.—A phenomenon of special interest arises when Newton's rings are produced with sodium light. As the thickness of the film increases, the clearness of the interference bands is alternately increased and diminished. This observation goes back to FIZEAU. The phenomenon depends on the fact that sodium light is not really homogeneous, but consists of two components which are very close together (p. 158), with wave-lengths 589·6 and 589·0 $m\mu$. Hence it is only in the region of the 500th ring that the one component of sodium light will give rise to a bright band where the other component gives rise to a dark band, while in the neighbourhood of the 1000th ring the two systems of rings will reinforce each other. The alternations of clearness and blurring are better observed with the wedge-shaped films produced, e.g. by laying one flat piece of plate glass on another in such a way that they touch along one edge and are separated by not too great an interval at the other, than with Newton's rings, which at a considerable distance from the point of contact move closer and closer together and hence can only be seen separately by using a very high magnification.

It has been found that even if the light used is as strictly monochromatic as possible, i.e. if one line is selected from a line spectrum and used alone, almost all sources of light exhibit a periodic variation of sharpness in interference bands where the difference of path is large. From this it is to be inferred that most of the lines in line spectra consist of components which lie very close together and cannot be resolved by means of prismatic spectrometers. Hence observations of interference phenomena with large differences of path have become an important means of investigating the "fine structure" of spectral lines. For some examples of this see p. 197 and Plate XV.

3. Curves of Equal Thickness.

Systems of interference bands which, like Newton's rings, are associated with regions where the thickness is the same, may be referred to as interference curves of equal thickness.

The superposition of interfering waves which gives rise to Newton's rings occurs in the region where the thickness is varying; hence the rings are not seen clearly unless the eye is accommodated for them.

The way in which such interference arises is illustrated once again in fig. 12. Two rays PA, PB start from the point source P. PA is reflected at the front surface of the plate, and reaches the retina of the observer's eye by the path

AEN. PB enters the plate at B, is then reflected at C by the rear surface, and leaves the plate at A, subsequently reaching the eye by the path ADN. As the eye is accommodated for the point A, the rays from A meet at a point N on the retina. The point N is bright or dark according as the rays are in the same phase or in opposite phases when they reach the retina. A difference of phase, if present, must of course have been present at the point A. As the rays leaving P are coherent, they are also in the same phase at both B and R, BR being perpendicular to PA. The difference of phase is accordingly determined by the two paths BC + CA and RA. The optical length of the path BC + CA is given by $\mu(\mathrm{BC} + \mathrm{CA})$. The difference of path between the two rays (Δw) is accordingly $\mu(\mathrm{BC} + \mathrm{CA}) - \mathrm{AR}$. If we assign the same value i to the angles of incidence of the two rays (which are very close together), it follows from fig. 13 that

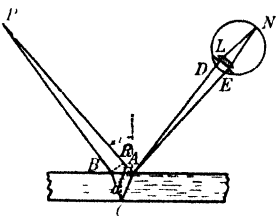


Fig. 12 — Curves of equal thickness observed by the eye

$$\Delta w = \mu(\mathrm{AC} + \mathrm{BC}) - \mathrm{AR} = 2\mu\mathrm{AC} - \mathrm{AR} = \frac{2\mu d}{\cos r} - \mathrm{AB}\sin i$$
$$= \frac{2\mu d}{\cos r} - 2\mathrm{AF}\sin i = \frac{2\mu d}{\cos r} - 2d\tan r \sin i$$
$$= \frac{2\mu d}{\cos r} - \frac{2\mu d \sin^2 r}{\cos r} = \frac{2\mu d}{\cos r}(1 - \sin^2 r) = 2\mu d \cos r$$
$$= 2d\sqrt{(\mu^2 - \mu^2 \sin^2 r)} = 2d\sqrt{(\mu^2 - \sin^2 i)}.$$

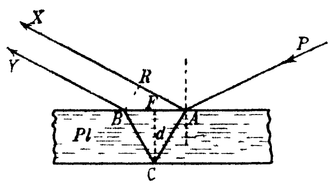


Fig. 13 — Difference of path between the interfering rays for curves of equal thickness

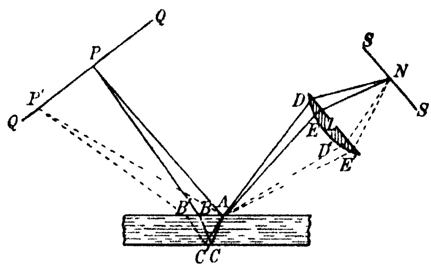


Fig. 14.—Curves of equal thickness projected on a screen

In the experiment the assumption that the two rays PA and PB (fig. 12) lie very close together is satisfied, for the reflected rays have to enter the small pupil of the observer's eye. The question of the form which the relationships take when this assumption is abandoned is important.

For this purpose we imagine the eye replaced by a screen SS with a condensing lens L in front of it (fig. 14). The two rays PA and PB starting from the point P of an extended source of light QQ (e.g. a spread-out bunsen flame coloured by common salt) behave in exactly the same way as the two corresponding rays in fig. 12. Then if a number of rays coming from A are to meet the lens L and be reunited at the point N, other points of the source of light, such as P′, must contribute to the effect. The rays from P′, whose path is shown dotted in fig. 14.

are incident on the plate at a different angle i'; hence their difference of path is $\Delta w' = 2d\sqrt{(\mu^2 - \sin^2 i')}$; for these two rays, therefore, the difference of path is not the same as for the rays emanating from P. Hence if the rays from P were to be extinguished, the rays from P′ might actually give rise to maximum intensity at N as a result of interference.

It follows from the above that with pencils of large angular aperture Newton's rings lose their sharpness.

4. Curves of Equal Inclination.

The expression for the difference of path between two rays which are caused to interfere by a plate, $\Delta w = 2d\sqrt{(\mu^2 - \sin^2 i)}$, involves i, the angle of incidence, as well as d, the thickness of the plate; that is, i also affects the interference curves. Newton's rings do actually move when one moves the eye while observing them, but this effect is of minor importance, as the effect of d greatly preponderates.

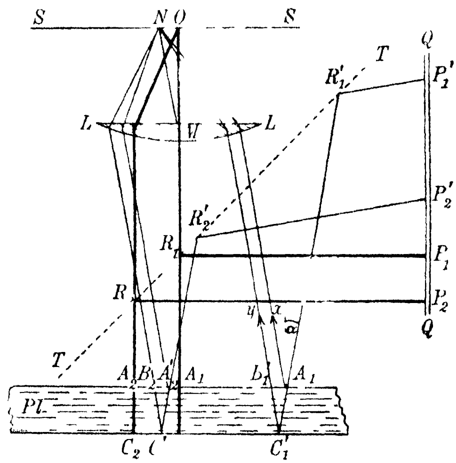


Fig. 15.—The production of curves of equal inclination

If, however, we use a plate which is bounded by very accurate parallel planes, so that d is invariable, interference curves can occur if the angle of incidence of the rays of light varies from point to point. The **curves of equal inclination** arising in this way were observed accidentally by Haidinger * in 1849, but he did not explain them or investigate them further; proper attention was first drawn to them by Lummer in 1884.

Unlike the curves of equal thickness, the curves of equal inclination are observed with the eye accommodated for infinity. Fig. 15 may serve to explain their mode of occurrence and to give a diagrammatic representation of apparatus by which they may be readily observed. Pl is a section of a plate with accurately flat parallel sides. A condensing lens LL is set up at a distance of about 50 cm. from it; in its focal plane is the screen SS, which is parallel to the plane of the plate. A transparent piece of plate glass TT is set up at an angle of about 45° between the plate Pl and the lens LL in such a way that an extended monochromatic source of light represented by the double

* W. von Haidinger (1795–1871), director of the *Geologische Reichsanstalt* in Vienna from 1849.

line QQ is made to illuminate the plate Pl. If the faces of the plate Pl are absolutely flat and parallel, a system of interference curves of equal inclination is produced on the screen SS; in the present case these are circles whose centre O lies on the perpendicular through M, the centre of the lens, to the plate and the screen. The rings of equal inclination may also be observed subjectively with the eye accommodated to infinity, for the latter has the same effect as the lens LL and the screen SS.

All the pencils of rays which leave the plate Pl at right angles to it meet at the point O of the screen SS. To each ray reflected at right angles to the front surface of the plate there corresponds a ray coherent with it, which is reflected from the rear surface of the plate, the difference of path between them being $2\mu d$. Thus at O there will be brightness or darkness according as this difference of path is an odd or even multiple of half a wave-length (there being a change of phase on reflection at the front surface). The corresponding pairs of rays reaching O all start from different points of the extended source of light. Thus, for example, the ray C_1A_1O comes from P_1, the ray C_2A_2O from P_2. At points of the screen SS on either side of O the rays reflected obliquely by the plate Pl are recombined in such a way that a parallel pencil inclined at a definite angle α is brought to a focus at every point of the screen, and in each of these pencils every ray reflected from the front surface of the plate is associated with a ray from the rear surface coherent with it, the difference of path being, as shown on p. 184 (fig. 13), $\Delta w = 2d\sqrt{(\mu^2 - \sin^2 i)}$. In fig. 15 two such pairs of rays, starting from the points P_1', P_2' of the source of light, are drawn for the point N. An interference pattern must accordingly be produced on the screen, in which a dark or more or less bright spot at O will be surrounded alternately by bright and dark rings, the transition from bright to dark being gradual.

As according to our assumption d is invariable, the difference of path Δw, and hence also the difference of phase of the components, depends only on the angle α.

The occurrence of *rings of equal inclination* is entirely independent of the aperture of the lens forming the images, and the source of light must be an extended one (contrast *rings of equal thickness*).

Rings of equal inclination can also be observed by looking through a plate with accurately flat and parallel surfaces at an extended monochromatic source of light, with the eye accommodated for infinity. The centre of the rings is then the foot of the perpendicular drawn from the eye to the plate. According to the phase difference of the two components leaving the plate at right angles, the centre appears dark or bright. If the plate is moved about between the eye and the source of light, the appearance of the centre will alter if the thickness varies slightly, a change from dark to bright always corresponding to a difference in thickness equal to a quarter of a wave-length. If the plate has accurately flat and parallel surfaces, Lummer's curves are complete circles. Conversely, if the circles are accurate, the accuracy of the plate may be inferred, as even slight variations in thickness manifest themselves by marked distortion of the circles. For this reason the curves of equal inclination furnish an extremely sensitive method for testing the flatness of plates.

5. Interference at Two Plates.*

If a piece of plate glass bounded by parallel planes is cut across the middle and the two parts P_1 and P_2 (fig. 16), inclined at a very small angle to one another, are set in a tube whose inner surface is blackened and which has a hole O in one of its ends, an observer B, looking at the bright sky through the tube, sees beside the hole O an image of the hole (A) which is crossed by dark and coloured interference bands running parallel to the geometrical line of intersection of the planes of the two plates P_1 and P_2. The greater the angle between the plates, the finer are the bands. The experiment will not succeed unless the two plates are cut from the same piece of glass and hence are of exactly the same thickness.

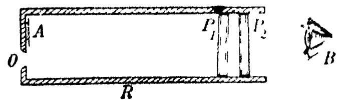


Fig. 16.—Brewster's interference band

In order to explain the phenomenon we begin by considering fig. 17, in which it is assumed that the two plates P_1 and P_2 are parallel to one another. At each surface the ray of light L is split up into a transmitted ray and a reflected ray. The transmitted ray is subject to refraction, but as this plays no essential part in what follows, it has been ignored in drawing the figure. Of the components of the incident ray which arise by repeated reflection, we consider only the four components 1, 2, 3, 4 shown in fig. 17, as these components can alone contribute to the formation of the lateral image in the experiment of fig. 16. The ray M which goes straight through produces the direct image of the hole and we shall therefore ignore it also.

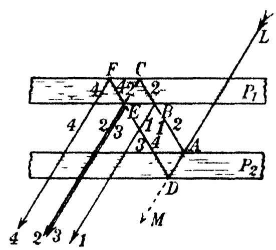


Fig. 17.—To illustrate the theory of interference at two plates

The ray 1 is reflected once at each of the inner adjacent surfaces of the plates P_1 and P_2. The ray 2 is reflected at the inner surface of P_2 and at the outer surface of P_1. The ray 3 is reflected at the outer surface of P_2 and at the inner surface of P_1. Finally, the ray 4 is reflected at the outer surfaces of both plates.

If the ray L falls on the plates obliquely, the two components 1 and 4 are far apart, although, to be sure, they will meet the other rays again on the retina of an eye accommodated for infinity. The two rays 2 and 3 coincide completely. If the two perfectly flat plates were set up exactly parallel to one another, the two components 2 and 3 would emerge without any difference of path. The two rays 1 and 4 differ considerably from 2 and 3 as regards length of path.

The same relationships are illustrated again in fig. 18, but with the difference that L is re resented as a point source of light and its images by reflection are constructed. Reflection at the faces A and D of the plate P_2 gives rise to the images L_A and L_D, and these give rise, by reflection at the faces B and C_1F of the plate P_1, to the four images L_1, L_2, L_3, and L_4, from which the rays corresponding to the four components mentioned above appear to start. Then if the plates P_1P_2 are parallel to one another and of exactly the same thickness,

* First observed in 1817 by D. Brewster (1781–1861), a noted Scottish physicist

the images L_2 and L_3 exactly coincide (although in the figure they are shown separately for the sake of clearness).

The relationships are altered, however, if the two plates are inclined at an angle to one another. Then, as is shown by fig. 19, the four images L_1, L_2, L_3, L_4 are again formed; the images L_2 and L_3, however, do not coincide, but are distinct, though lying close together. As the rays producing these images are coherent (since they come from the same source L), the light rays starting from the images must give rise to interference bands in exactly the same way as the two images by reflection do in Fresnel's mirror experiment. The rays 2 and 3

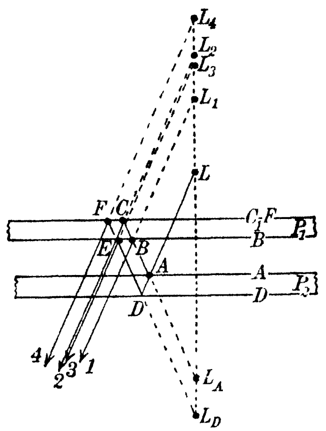


Fig. 18.—Reflection at parallel plates

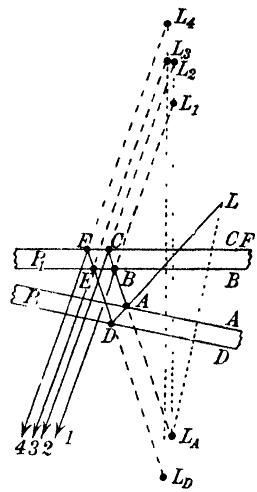


Fig. 19.—Reflection at plates inclined at an angle to one another

which (apparently) come from L_2 and L_3 are parallel. Hence if they enter an eye accommodated for infinity and are reunited at a point on the retina, the eye must see the interference bands directly. If we imagine that there are one or more other point sources of light beside L, their images will occupy similar positions to L_2 and L_3, and these rays produce at the same point of the retina of the eye accommodated for infinity exactly the same interference phenomena as the images by reflection of the point L which we considered previously. For this reason the hole O in the tube shown in fig. 16 may be fairly large, but not so large that the interference image A, which is seen at infinity since the eye is focused for infinity, coincides with the hole, for then the interference pattern would be masked by the light from the bright opening.

A variant of this experiment is shown diagrammatically in fig. 20. In this apparatus, which is due to JAMIN* (1858), the two interfering rays 2 and 3 are separated from one another in space over a part of their path. If a transparent body of any kind is inserted in the path of one of the rays, and its refractive index differs from that of air, the wave-length of the light, and hence the position

* J. C. JAMIN (1818–86), a French physicist.

of the interference bands, is altered. By means of Jamin's apparatus it is possible e.g. to demonstrate and measure the variation of the refractive index of air with the temperature or with the pressure very accurately. It is extensively used in practice, e.g. in fuel research for testing the composition of the products of combustion, and in the Zeppelin airship for testing the hydrogen content of the air in the body of the airship.

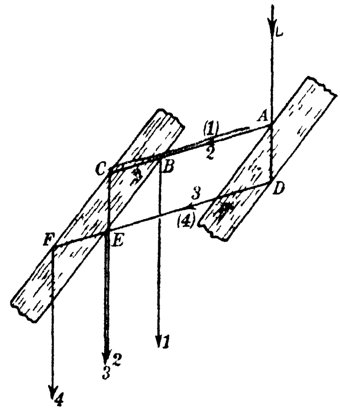


Fig 20 —Diagram illustrating Jamin's interferometer

The actual lay-out of JAMIN's refractometer is illustrated in fig. 21. Two glass tubes T, one containing air and the other the gas under investigation, are placed between the two thick glass plates P. The interference bands are observed through the eyepiece E. C is a compensator consisting of two thin glass plates which can be tilted so as to retard one of the light rays relative to the other by any desired amount.

Both Brewster's apparatus and Jamin's apparatus are very well

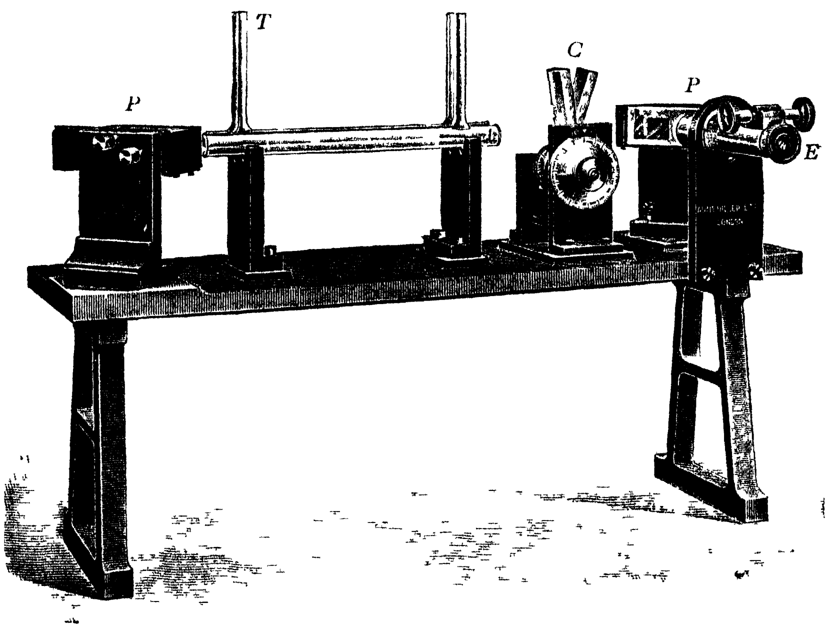


Fig. 21.—Jamin's refractometer (Adam Hilger, Ltd.)

adapted for the objective demonstration of interference bands, as the source of light may be an extended one and hence may be made very intense.

6. Michelson's Interferometer.

One of the most interesting types of apparatus for producing interference is MICHELSON'S* interferometer, one of the commonest forms of which is shown diagrammatically in fig. 22. A ray of light S falls at an angle of 45° on a plane semi-transparent silvered glass plate P and is there split up into the transmitted ray 1 and the reflected ray 2. The ray 1 falls on the mirror Sp_1 and the ray 2 on the mirror Sp_2, in both cases normally, hence each ray is reflected into itself (in fig. 22 the two rays, going and returning, are drawn separately for the sake of clearness). On their return journey the two rays fall

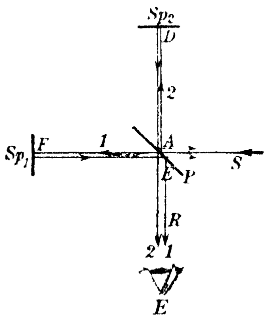


Fig. 22.—Michelson's experiment (diagrammatic)

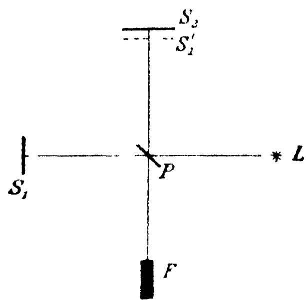


Fig. 23 --Rough diagram illustrating the Michelson interferometer

on the glass plate P again and each of them is again split up into two parts. Of these, however, we shall consider only the parts shown in the figure by continuous lines; the two other parts, which are indicated by dotted lines, return to the source of light.

If the distances of the two mirrors Sp_1 and Sp_2 from the plate P (whose thickness we shall in the first instance neglect) are equal, the two rays 1 and 2 have travelled over the same length of path, that is, they emerge in the direction R without any difference of phase and reinforce one another. Reinforcement also occurs when the distances of the mirrors from the glass plate differ by an even multiple of a quarter wave-length, i.e. when the lengths of path of the light rays differ by an even multiple of half a wave-length. On the other hand, the two rays cancel one another when the difference of the distances of the two mirrors from the glass plate is an odd multiple of a quarter wave-length. To an eye E focused for infinity, on which

* ALBERT ABRAHAM MICHELSON, born at Strelno (Posen), was professor at Chicago, received the Nobel prize in 1907, died in 1931 (see also pp. 201, 209, 217).

the rays 1 and 2 fall, the field of view accordingly appears alternately bright and dark when one of the mirrors Sp_1, Sp_2 is moved, and in fact a bright region becomes dark when one of the mirrors is moved through a quarter wave-length. It follows that this apparatus may be used for the absolute measurement of wave-lengths if the mirror Sp_2, say, is made to move parallel to itself along the direction of the light ray

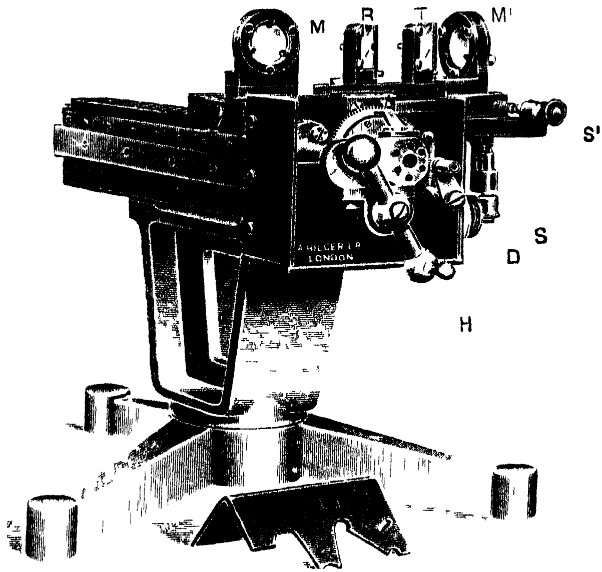


Fig 24 —Michelson's interferometer (Adam Hilger, Ltd)

through a distance which can be measured by a micrometer screw, and the number of alternations of light and darkness in the field of view is observed at the same time.

The field of view of course does not appear uniformly bright or dark but (if the whole apparatus is accurately centred) exhibits concentric circles; for the source of light S is an extended one, so that a number of parallel pencils of rays fall at different angles on the glass plate P. Circular curves of equal inclination are accordingly formed. This, however, is of minor importance as regards the carrying out of the measurements, as of course it is only necessary to observe an arbitrary point, e.g. the centre, of the field of view.

The way in which the Michelson interferometer works may be readily grasped by imagining one arm of the apparatus rotated towards the other (S_1 towards S_1' in fig. 23). If the planes of the two mirrors are parallel, the interference pheno-

mena described above may immediately be deduced from those produced by a plate bounded by parallel planes.

An actual instrument is shown in fig. 24. The mirror M is mounted on a travelling carriage actuated by the handle H, while M′ is a fixed mirror, which, however, may be adjusted by means of the screws S and S′; R is the plate at which the rays are split up, and T a compensating plate (see below).

In spite of the very great experimental difficulties MICHELSON succeeded, first in collaboration with MORLEY and then with BENOÎT at Breteuil, in evaluating the length of the standard metre (Vol. I, p. 7) in wave-lengths of light.

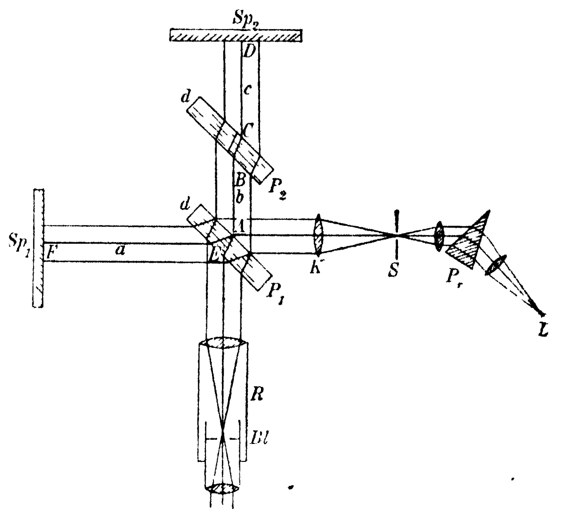


Fig 25.—Michelson's experiment; the complete apparatus

A more detailed diagram of the experiment is given in fig. 25. As the source of light MICHELSON used electric sparks between cadmium points. The cadmium light was then decomposed by the spectroscope P_r, and the red cadmium line was selected by means of the slit S. The light from this was rendered parallel by the collimating lens K and fell on the flat glass plate P_1. Here the pencil of parallel rays was split up as described above and reflected by the two mirrors Sp_1 and Sp_2 in such a way that the components emerging in the direction R entered a telescope R focused for infinity. The interference pattern was then produced in the plane of the stop Bl in the telescope.

The purpose of the flat plate P_2 is to compensate for the passage of the rays through the plate P_1. The difference of optical path between the two rays which interfere is accordingly equal to the difference between the lengths of the two paths in air, i.e. $\Delta w = 2(a - (b + c))$.

The experiment essentially consists in counting the number of alternations of light and darkness which occur when one mirror is moved

through a definite distance. The details of the apparatus are fairly complicated, and in order to evaluate the metre in terms of the wave-length of cadmium light intermediate steps must be inserted, as it is not possible to make waves with a difference of path of a metre interfere with one another. The results of the experiments give

$$\begin{aligned} 1 \text{ metre} &= 1{,}553{,}164{\cdot}1\ \lambda_R \\ &= 1{,}966{,}249{\cdot}7\ \lambda_G \\ &= 2{,}083{,}372{\cdot}1\ \lambda_B, \end{aligned}$$

where λ_R, λ_G, λ_B are the wave-lengths of the red, green, and blue cadmium lines respectively, these values being for dry air at 760 mm. pressure and at 15° C. on the hydrogen scale. These measurements are probably the most accurate which have ever been carried out, as their accuracy reaches to the eighth significant figure.

There is another possible method of producing interference bands. If the plane of one mirror makes an angle with that of the other when the arms are imagined to be superposed as described above, the interference pattern of a wedge-shaped plate is formed. Thus, e.g. if one plate is rotated slightly relative to the other about a vertical axis (fig. 26), we obtain a family of straight interference bands running from top to bottom; the greater the inclination of the plates to one another, the smaller the distance between the bands. If, for example, the distances of the centres of the mirrors from the separating plate are equal, the difference of path is zero, and using white light we see the central interference band with the coloured bands grouped almost symmetrically about it. Then if one arm is lengthened or shortened slightly in any way, the central position relative to the arms is no longer at the centre of the mirror but to the left or right of it, as we may see from fig. 26, in which the new position is shown by dots. That is, the central band now appears at a different part of the mirror. If, therefore, the inclination of the planes of the mirrors to one another is such that about three bands are to be seen on either side of the central band (that is, after careful rotation of one arm in the direction of the other the left-hand end of one mirror is six wave-lengths nearer than the right-hand end), the displacement of the central band through a distance equal to the breadth of a band means that the length of one arm has been altered by one wave-length. As it is not difficult to produce bands about 5 mm. across, this means that a change of length is multiplied by 10^4. By observation through a telescope, and more particularly by photographic recording, it is possible to detect shifts of the bands amounting to only a thousandth of their breadth, i.e. a change in length of 10^{-7} cm. (and this under certain circumstances for a length of many metres, corresponding to a relative change in length of 10^{-11}). For an application of this method see Vol. V (MICHELSON'S interference experiment).

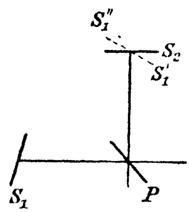


Fig. 26 —The interferometer with the mirrors rotated through an angle

7. Interferometry.

In order to test the homogeneity of single spectral lines accurately, we make the light from them give rise to interference with a very

large difference of path (p. 183). This method reveals "fine structure" in lines which no ordinary spectroscope is capable of resolving into components. The method is known as interferometry; it goes back to FIZEAU (1862) (p. 183), but was first developed for general use by A. A. MICHELSON (1892).

In all the experiments hitherto described in which interference was brought about by inserting a plate in the path of the rays, we have assumed that it is only the rays reflected *once* at the front surface and *once* at the rear surface of the plate that give rise to interference, when the interference bands are observed by reflected light (cf. figs. 13–15, pp. 184, 185). In the case of the interference of the transmitted light, we only consider the ray which goes straight through and the ray which is transmitted after being reflected *once* at the rear surface and *once* at the front surface of the plate. On the other hand, we have ignored the components which are split off by repeated reflection at the surfaces of the plates. In actual fact these simplifications have hitherto been justifiable, as the intensity of the latter components was vanishingly small compared with that of the two first components.

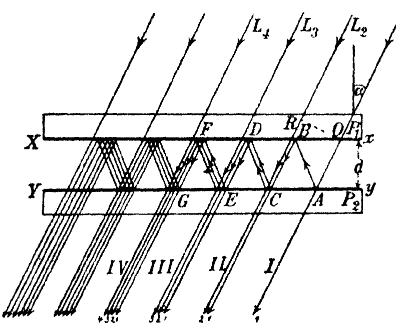


Fig. 27.—Interference when the differences of path are very great

The relationships, however, are essentially different when the arrangements are such that the subsequent components have an intensity which is still finite when compared with the intensity of the two first components.

The Fabry-Pérot Interferometer (1897).—Two slightly wedge-shaped glass plates (the angle of the wedge being about half a degree) are lightly silvered on one side, in such a way that part of the light is transmitted by the layer of silver, but the reflected portion contains a large percentage of the total energy of the light. The two glass plates are placed with their silvered faces next one another in such a way that a layer of air bounded by accurately parallel planes is formed between the layers of silver. The apparatus is shown diagrammatically in fig. 27; P_1, P_2 denote the two glass plates with the adjacent silvered faces Xx, Yy, which enclose between them a layer of air of thickness *d*.

If a pencil of coherent parallel rays L_1, L_2, L_3, . . . falls on the system of plates, the angle of incidence being α, part of each ray is reflected at the layer of silver Xx and so far as we are concerned here is lost and hence of no account. We consider only that portion of the light which enters the layer of air. (The refraction of the ray of light in the glass is not shown, as it is of no importance for us.) The portion of the ray L_1 entering the air at O is partly reflected at A by the second layer of silver and partly transmitted through the plate P_2 without change of direction. The reflected ray meets the first silver layer Xx at B and is

then re-reflected to C on the second silver layer Yy. It is thus repeatedly reflected along the path OABCDEFG. . . . At each point of reflection, part of the light travels outwards. In the figure only those portions are shown which pass outwards through the second plate, as the others do not interest us. We see that each successive portion of the ray L_1 travels over a path in the layer of air which exceeds the length of the path of the preceding one by the amount $\Delta w = OA + AB - BR$ (where BR is to be imagined as measured in air). This expression may be transformed into

$$\Delta w = 2d \cos\alpha.$$

If instead of the layer of air we had a plate of a substance with refractive index μ, the difference of path would be given by the expression

$$\Delta w = 2d\sqrt{(\mu^2 - \sin^2\alpha)}$$

which we calculated previously (p. 184).

What we have deduced for the portions of the ray L_1 holds in exactly the same way for the portions of the other parallel rays L_2, L_3, . . . As all the rays are coherent, so are the portions of them, so that the emerging rays are capable of interfering with one another. If we now imagine a convergent lens placed in the path of the rays, and a screen behind it parallel to the layer of air and at a distance equal to its focal length, all the parallel components will be made to converge by the lens and the interference diagram of all the rays will thus be collected at a point. This point Q lies in a line making an angle α with the axis of the lens (fig. 28).

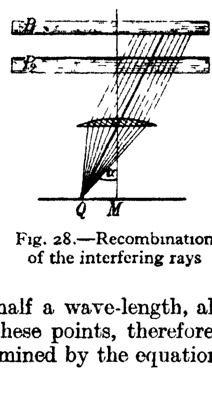


Fig. 28.—Recombination of the interfering rays

The brightness of the point Q depends on the relationship between the wave-length λ and the difference of path Δw. To begin with we assume that the light used is perfectly monochromatic. It is then clear that at all points Q_b for which Δw is an integral multiple of a wave-length, or, what is the same thing, an even multiple of half a wave-length, all the components will arrive at Q in the same phase. At these points, therefore, there will be maxima of intensity; their positions are determined by the equation

$$2d \cos\alpha = 2k \cdot \frac{\lambda}{2}.$$

Further, there are minima of intensity at those points Q_d where the difference of path Δw is equal to an odd multiple of half a wave-length, as at these points the component (say of the ray L_1 (fig. 27)) produced by a single reflection at the two silver surfaces arrives in a phase exactly opposite to that of the component of the next ray (L_2) which travels straight through. As, however, the intensities of the two components are not exactly equal, the illumination is merely weakened and not destroyed. The components of L_1 split off by double, triple, and multiple reflection meet the direct component of L_3 in the same phase and that of L_4 in the opposite phase, and so on. In any case a great deal of the light is cut out and the points Q_d exhibit minima of intensity, i.e. darkness. The positions of the intensity minima are determined by the equation

$$2d \cos\alpha = (2k + 1)\frac{\lambda}{2}.$$

At the points lying between the points of maximum brightness and the points of maximum darkness the various components again partially extinguish one another. For if the difference of path between the directly transmitted

component and the component split off by a *double* passage to and fro between the layers of silver is equal to an odd multiple of half a wave-length, these two portions have opposite phases, i.e. partially destroy one another. The same occurs for points such that the difference of path between the directly transmitted component and the component split off by a *triple* passage between the two layers of silver, is equal to an odd multiple of half a wave-length, i.e. in general when the difference of path after *repeated* passage between the two layers of silver is equal to an odd multiple of half a wave-length.

It follows that it is only at points for which a is given by the equation

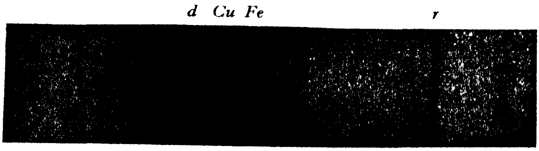
$$2d \cos a = 2k \cdot \frac{\lambda}{2}$$

or

$$\cos a = k \cdot \frac{\lambda}{2d}$$

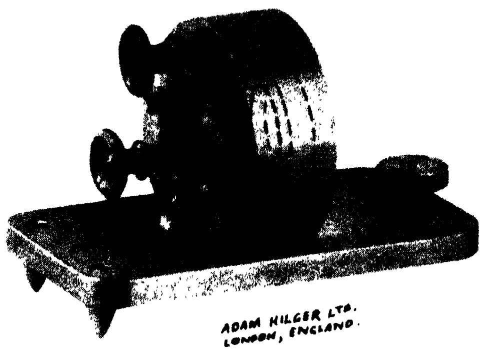
that all the components reinforce one another. These must stand out against the background as particularly bright points. If the source of light from which the rays start is a point source, sharply defined concentric circles are formed on a dark background about a bright central point.

If we now use a source of light which emits light of two wave-lengths differing only slightly from one another, each gives rise to its own system of circles, just as we found to happen in the case of Newton's rings when they are produced by the two components of sodium light (p. 183). Here, however, the following distinction is to be noted. Newton's rings are produced by the interference of *two* pencils of rays, whereas the circles occurring with the present apparatus are due to the interference of a *large* number of coherent pencils or rays (possibly exceeding twenty). As is shown in detail in the following section, this has the result that in the first case the maxima are approximately as broad as the minima (fig. 41, p. 203), whereas in the second case, although the maxima occupy the same positions as they do when two pencils interfere, they are much brighter, and much narrower than the minima lying between them (fig. 42, p. 204). In the first case, therefore, the circles become blurred when a circle of one system falls in the space between two circles of the other system, while in the present case the two systems of circles, owing to their sharpness and their rapid fall in intensity, remain entirely separate when a circle of one system falls between two of the other system. If the wave-lengths are λ_1 and λ_2 and the corresponding angles of incidence a_1 and a_2, the difference $a_2 - a_1$, which is a measure of the distance between the maxima, must increase with k, the order of the difference of path. Hence the maxima for the two kinds of light move farther apart, and hence are easier to observe separately the greater the order k is.

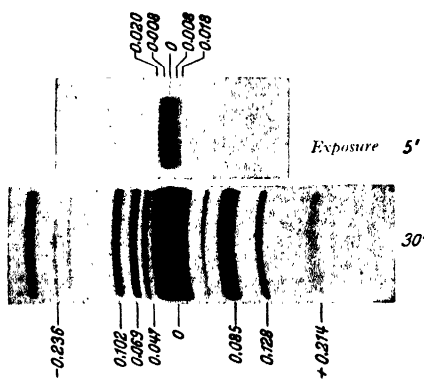


Ch. VII, Fig. 20.—Prismatic spectrum of X-rays obtained with the apparatus of fig. 19 (p. 165). Fe $K\alpha$ = 1·933 Å.; Cu $K\alpha$ = 1·538 Å.; $\delta = 1 - \mu = 12{\cdot}38 \cdot 10^{-6}, 8{\cdot}125 \cdot 10^{-6}$ respectively (r reflected ray, d direct ray). (From *Handbuch der Experimentalphysik*, Vol. 24, Part II (Akademische Verlagsgesellschaft, Leipzig).)

Ch. VIII, Fig. 29.—Fabry-Pérot interferometer (etalon) (Adam Hilger, Ltd.).



Ch. VIII, Fig. 30*a*.—Fine structure of the blue helium line λ = 4713 Å., obtained with Fabry-Perot interferometer (the distance between the plates being 17·1 mm.). The three components of this line are related as shown by the three strokes. Time of exposure, 2 hours 40 minutes. Distance between the components, 0·017 and 0·36 Ångstrom units respectively. (Photograph by Hansen.)



Ch. VIII, Fig. 30*b*.—Fine structure of the green mercury line 5461 Å. Numerical values are in Ångström units. Lines without a number attached belong to neighbouring orders. Note that on the same scale the distance between the two D lines of sodium would come to 47 cm. (Photograph by Lau; from *Zeitschrift fur Physik*, Vol. 63 (Springer, Berlin).)

Fig. 29 (Plate XV) shows a practical form of the apparatus. The distance between the glass plates is determined by invar rings. Fig. 30*a* (Plate XV) gives the photograph of a helium line; parts of the ring systems of the various orders are clearly seen. This line turns out to be a triplet, i.e. a group of three lines. To obtain the sharpness necessary for such investigations, the discharge tube containing helium had to be cooled by liquid hydrogen. By coupling two such pieces of apparatus in series a still greater resolving power has been attained (multiple interferometer). Fig. 30*b* (Plate XV) gives an idea of the resolving power which it is possible to reach (note the scale).

The Lummer-Gehrcke Plate.—In collaboration with Gehrcke,* Lummer (1902) brought forward an interferometer in which the multiple reflection of a pencil of rays in the interior of a glass plate bounded by parallel planes is used to produce interference with a large difference of path. The essential feature of this apparatus is a perfectly flat glass plate P with parallel sides, to one end of which a small right-angled prism Pr is cemented with Canada balsam (fig. 31).

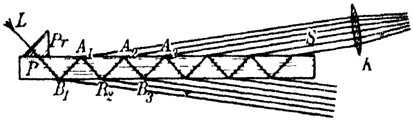


Fig. 31.—Lummer-Gehrcke plate

If a pencil of rays L passes through the prism Pr into the glass plate P as shown in the figure, it meets the plane boundary surfaces at an angle which is nearly equal to the critical angle of total reflection; hence parallel rays emerge from the glass plate, almost grazing the surface, at the points A_1, A_2, A_3, . . . , and B_1, B_2, B_3, . . . on either side. Just as before (p. 184), the difference of path is given by $\Delta w = 2d\sqrt{(\mu^2 - \sin^2 a)}$, where d is the thickness of the plate, μ its refractive index, and a the angle of incidence measured in air. All the pencils S emerging from one side surface are then reunited by the convergent lens K and made to interfere. K is the objective of a telescope.

The theory of the Lummer plate agrees almost entirely with that of the Fabry-Pérot apparatus; the device of making the repeated reflections inside the glass plate take place almost at the critical angle of total reflection has the effect that there is only a very trifling loss of intensity at each reflection, so that the number of reflections may be made extremely large. Thus a very large number of parallel rays are made to interfere, and hence the distribution of intensity in the interference pattern exhibits unusually sharp maxima, consisting of extremely fine sharp lines on an almost completely dark background.

* E. Gehrcke, professor at the *Physikalisch-Technische Reichsanstalt*, Berlin.

8. Interference of Parallel Rays arising from Diffraction of Parallel Rays.

Fraunhofer Diffraction Phenomena.—The phenomena investigated in Chap. I, § 6 (p. 11) and § 1 of the present chapter (p. 174), in which diffracted rays of light act on a point at a *finite* distance, are referred to as Fresnel diffraction phenomena. The relationships become simpler if we investigate the interference arising from light rays which are still *parallel* after diffraction. We then have the so-called Fraunhofer diffraction phenomena. In order that these rays may be made to interfere at a finite distance, they must be made to converge at a point by means of a convex lens or concave mirror. According to Chap. V, § 5 (p. 102), no change of phase is produced by the interposition of lenses or mirrors in the path of the rays; hence we can discuss the phenomena in exactly the same way as if the parallel rays interfered directly.

The apparatus which may be used to produce the phenomena discussed below is shown diagrammatically in fig. 32; L is a point source of light, which in conjunction with the collimating lens B gives rise to the pencil of parallel rays S. Instead of the point source L, we may very well use a slit brightly illuminated from behind. If this is sufficiently far away the rays from it may be regarded as parallel even in the absence of a collimating lens. In the path of the rays we also place the condensing lens K, which (if the light rays are not otherwise modified in any way) will produce a real image O of the source L on the screen Sc in the focal plane of the lens. If, however, a body giving rise to diffraction (denoted by GG in the figure) is inserted anywhere into the path of the rays, each of its points will send out rays in all directions. Each pencil of parallel rays is now brought together again by the lens L at a point A on the screen determined by α, the inclination of the rays. The two lenses B and K may be replaced by a single one, thus simplifying the apparatus. In our discussion, however, we shall ignore this possibility and shall assume that the diffracting body GG is illuminated by parallel light. The condensing lens K may be replaced by the objective of a telescope. The focal plane common to the objective and the eyepiece of the telescope then corresponds to the screen.

In future we shall specially consider the path of the rays from the diffracting body GG to the lens, or, what is the same thing, we assume that the parallel rays starting from the diffracting body meet and interfere at infinity.

Diffraction at a Single Slit.—Let the diffracting body denoted by GG in fig. 32 be a single narrow slit. At the centre O of the screen we then observe a bright image of the source L, which is accompanied on either side by a large number of parallel bright and dark bands of diminishing intensity (cf. fig. 38. Plate XVI).

If, following Huygens' principle, we regard each separate point of the opening of the slit as the starting-point of a new system of waves, we in the first place obtain a *maximum* of intensity where the perpendicular to the slit at its centre meets the screen; for here all the incident rays are in the same phase.

To explain the two first *minima* which lie on either side of the central maximum, we imagine that the series of points XY which form the slit (cf. fig. 33,

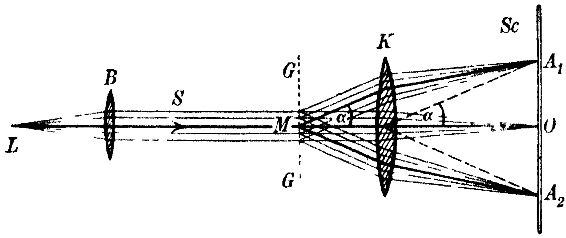


Fig 32 —Fraunhofer interference of parallel light

which represents a section perpendicular to the length of the slit) is divided into *two* parts at the centre M. We shall now assume that the number of separate centres of excitation in the slit opening is 100; this number is quite arbitrary, as we might equally well take it as 1000 or 100,000 (fig. 33 is drawn for 24). A dark band is then formed on the screen behind the slit at the place where the waves from the first centre (X) and the fifty-first centre (M) cancel one another; for when the difference of the distances of these points from the point on the screen

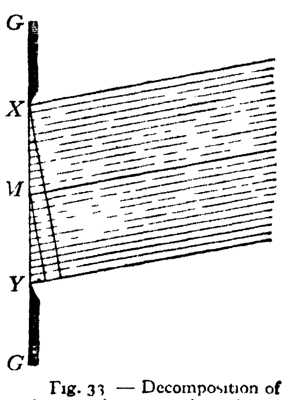


Fig. 33 — Decomposition of the pencil passing through the slit for the first minimum

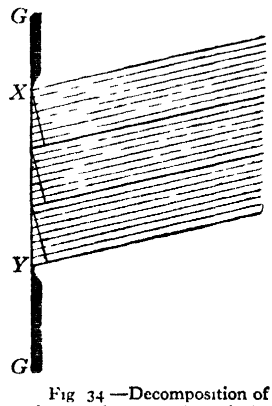


Fig 34 —Decomposition of the pencil passing through the slit for the first maximum

under consideration amounts to half a wave-length, the difference between the distances of the point from the second and fifty-second centres, the third and fifty-third centres, &c., the fiftieth and hundredth centres is also half a wave-length; in other words, the first minimum of intensity or first dark band occurs where the outermost rays of the pencil of parallel rays under consideration have a difference of path of one wave-length.

To explain the first *maxima*, i.e. the lateral bright bands following on the dark bands, we imagine the series of points forming the slit divided into *three*

equal parts (fig. 34). The position of the first minimum is such that there is a phase difference of half a wave-length between the rays starting from each point of the first third of the series of points and the rays from the corresponding points of the second portion, so that these rays completely destroy one another; on the other hand, all the light coming from the final third of the slit remains unextinguished. The difference of path between the outermost rays of the diffracted pencil of parallel light is then $3\lambda/2$.

For the next *minimum* we must divide up the breadth of the slit into four equal parts; the rays from these will completely cancel one another in pairs. The difference of path between the outermost rays of the diffracted pencil of parallel light is then 2λ.

In fig. 35 XY is the diffracting slit. Let the pencil of rays indicated by nine parallel lines, the central being MA, make the angle α with the axis MO. We now draw XZ from X perpendicular to the pencil of rays; let it cut off the segment $YZ = n\lambda/2$ from the boundary ray passing through Y.

If we now put $XY = d$, it follows immediately from fig. 35 that

$$\sin\alpha = \frac{n\frac{\lambda}{2}}{d}.$$

According to the above discussion, the pencil of rays gives rise to maxima and minima of intensity at all points of the screen for which n is a whole number (except $n = 1$).

Minima of intensity occur for even values of n, *maxima of intensity for odd values of* n.

The maxima, however, decrease in intensity as n increases, for only the nth part of the pencil of rays is effective in producing them. Hence the maxima and minima gradually blend with one another. In fig. 36 the intensities are plotted as ordinates and used to construct an intensity curve (cf. fig. 35, p. 145).

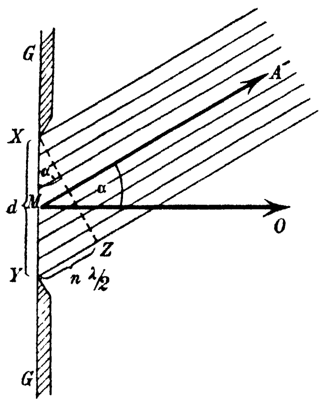


Fig. 35.—The outer rays of the pencil passing through the slit have a difference of path of $n\lambda/2$.

Diffraction by a Double Slit.—If the diffracting body (fig. 37) consists of a screen with two parallel slits of breadth d, X_1Y_1 and X_2Y_2, whose corresponding edges X_1, X_2 and Y_1, Y_2 are a distance b apart, each slit gives rise separately to a diffraction pattern identical with that described before.

In addition to this, however, the two parallel pencils S_1 and S_2 interfere with one another.

From X_1 and Y_1 we draw the perpendiculars to the direction of the rays and assume that the perpendicular from X_1 on the ray from X_2 cuts off the segment

$X_2Z_1 = k\lambda/2$. The angle β at which the pencil of rays is inclined to the axis MO is determined by the equation

$$\sin\beta = \frac{k\frac{\lambda}{2}}{b}$$

derived from the triangle $X_1X_2Z_1$.

If k is an *odd* number, the corresponding rays have exactly opposite phases at X_1 and Z_1; hence if a convergent lens (K in fig. 32, p. 199)

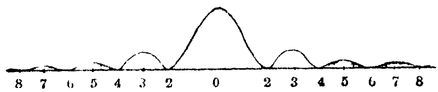


Fig. 36.—Maxima and minima of Fraunhofer diffraction bands due to a slit

is interposed, they cancel one another completely. But all the other rays of the pencil S_1 are in exactly the opposite phase to the corresponding rays of the pencil S_2; hence all the rays of the two pencils cancel in pairs, giving rise to a *minimum* of intensity.

If, on the other hand, k is even, all the rays of the pencil S_1 are in the same phase as the corresponding rays of the pencil S_2 and when brought together again by a convergent lens give rise to a *maximum* of intensity. The amplitude of vibration is doubled and the intensity quadrupled.

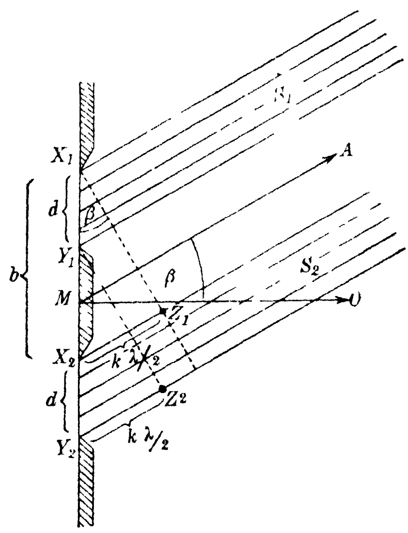


Fig. 37.—Diffraction due to two slits

The interference pattern due to the joint effect of the *two* slits is of course superposed on the interference pattern due to each slit by itself; hence we find that the diffraction pattern discussed previously is crossed by a number of bright and dark bands which are due to the joint effect of the two slits (fig. 38, Plate XVI).

Michelson's Method for Determining the Angular Separation of Double Stars and Stellar Diameters.—The centre of the interference pattern of the double slit corresponds to the image of the source on the screen as calculated by geometrical optics. If beside L we set up another point source in a direction at right angles to the line of the slit, it will give rise to an exactly similar system of interference bands, superposed on the system due to the first source. If the displacement of the second system relative to the first is exactly such that the maximum of intensity in the first system falls on the minimum of intensity in the second

system, and the two sources are of the same brightness, the surface of the screen appears uniformly bright, as the interference phenomena disappear. Here, provided that the angles are very small, i.e. the sine may be replaced by the arc, we have $\delta = \lambda/2b$, where δ denotes the angular separation of the two sources of light.

In this way it is possible to determine the direction of the line joining the two components of a double star and their angular separation, provided they do not differ too greatly in brightness. Two slits which can be moved relative to one another are mounted over the objective of a telescope, and the line joining the slits is rotated about the axis of the telescope and the distance between the slits is altered until the interference bands vanish. By increasing the distance b (up to 17 metres) the resolving power may be increased still further; we thus obtain the apparatus shown in fig. 39. An extended source of light (e.g. a small circular disc) will similarly cause the interference bands to disappear when the distance between the slits reaches a certain value, owing to the superposition of the interference systems arising from the various points of its surface; the angular diameter of the disc may again be determined (although we shall not go into further details here) from the distance between the slits for which the interference bands vanish. In this way ANDERSON first was able to resolve some double stars in 1919 and MICHELSON and PEASE were able to determine the diameter of some fixed stars (Vol. V). Attempts have recently been made to apply this method also to determine the magnitude of very small particles under the microscope.

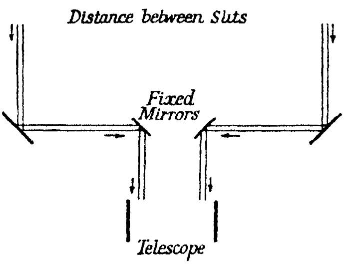


Fig. 39.—Michelson's method of determining stellar diameters

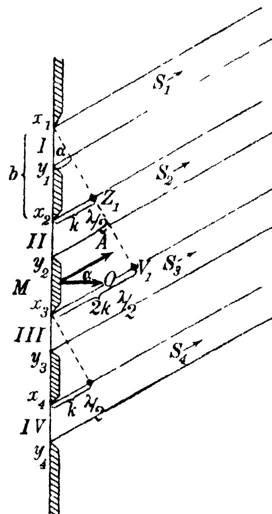


Fig. 40.—Diffraction due to four slits

Diffraction by Four Slits.—Of the four pencils of parallel rays S_1, S_2, S_3, S_4 diffracted by the slits I, II, III, IV (fig. 40), S_1 and S_2 give rise to interference bands identical with those explained above. An exactly similar system of interference bands is produced by the two pencils S_3 and S_4. The effects of these two pairs of pencils are accordingly additive. Maxima of intensity occur where the difference of path of the corresponding rays of the two pencils S_1, S_2 and S_3, S_4 is an even multiple of half a wave-length; that is, for the maxima of intensity the quantity k in the expression $\sin\alpha = k\frac{\lambda}{2}/b$ must be an even number. Where k is an odd number the light will be completely extinguished.

Tabulating the values of $\sin\alpha$, we obtain

$$\text{Maximum brightness for } \sin\alpha = 0,\ 2\cdot\frac{\lambda}{2b},\ 4\cdot\frac{\lambda}{2b},\ \ldots;$$

$$\text{Darkness for } \sin\alpha = 1\cdot\frac{\lambda}{2b},\ 3\cdot\frac{\lambda}{2b},\ 5\cdot\frac{\lambda}{2b},\ \ldots.$$

In addition to this, however, the pencil S_1 interferes with S_3 and the pencil S_2 with S_4. By fig. 40 these pencils have a difference of path of $2k\cdot\frac{\lambda}{2}$, as the perpendicular from x_1 on the corresponding ray of the pencil S_3 cuts off the segment $2k\cdot\frac{\lambda}{2}$. The interference of these pairs of parallel pencils accordingly gives rise to maxima of intensity at those points where $2k\cdot\frac{\lambda}{2}$ is an even multiple of half a wave-length. That is, for maximum brightness $2k$ must be an even number; hence for the maxima k must have the values 0, 2/2, 4/2, 6/2, ... On the other hand, the pairs of rays S_1 and S_3, S_2 and S_4 cancel one another where $2k$ is an odd number; hence for the dark regions k must have the values 1/2, 3/2, 5/2, 7/2, ... Hence we have

$$\text{Maximum brightness for } \sin\alpha = 0,\ 2\cdot\frac{\lambda}{4b},\ 4\cdot\frac{\lambda}{4b},\ 6\cdot\frac{\lambda}{4b},\ 8\cdot\frac{\lambda}{4b},\ \ldots;$$

$$\text{Darkness for } \sin\alpha = 1\cdot\frac{\lambda}{4b},\ 3\cdot\frac{\lambda}{4b},\ 5\cdot\frac{\lambda}{4b},\ 7\cdot\frac{\lambda}{4b},\ \ldots.$$

In fig. 41 the positions of the maxima are indicated by small circles and those of the dark regions by black spots. The points marked on the straight line 2 indicate the distribution subject to the assumption that there are only two slits or that if there are more slits they do not interfere with one another

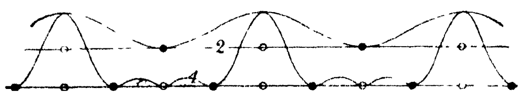


Fig. 41.—Intensities of the maxima and minima in diffraction by four slits

except for adjacent pairs. The points marked on the straight line 4 mark the positions of the bright and dark regions with the assumption that when there are a number of slits one pencil always interferes with the next but one, i.e. the pencils act like S_1, S_3 or S_2, S_4 in the above discussion.

Now it is clear that where two neighbouring pencils of rays extinguish one another there will be darkness under all circumstances, and that this fact cannot be affected at all by combining the pencils of rays in any other manner. Hence to determine the total effect of the four slits we have the following facts: there must be darkness in all circumstances wherever neighbouring slits produce dark interference bands, even when a pencil of rays would give maximum brightness if it were combined with the next but one. (For the two maxima resulting from this combination are of opposite phase, i.e. cancel one another when superposed.) On the other hand, a maximum produced by adjacent slits must be reinforced if the slits taken alternately or in some other definite way give rise to a maximum at the same place.

Above the straight line 2 in fig. 41 there is drawn a curve which represents the distribution of light in the interference phenomenon arising from two slits

or from pairs of neighbouring slits. The curve of course meets the straight line 2, which serves as axis of abscissæ, at the dark points and attains its greatest height above the circles.

Above the straight line 4 we then draw the intensity curve on the assumption that the slits also act alternately, i.e. that one slit is always combined with the next but one. This curve meets the axis of abscissæ (4) both at the points where there is a black spot on the line 2 and where there is a black spot on the line 4. At the points where there is a circle on both curves the curve rises to a height above the line 4 which is twice that of the first curve above the line 2. The latter curve shows that the intensified maximum rises steeply above the remaining points of the line 4.

In the figure the decrease in intensity of the maxima to either side has not been taken into account. In actual fact the simple intensity curve is not a sine curve but is as shown in fig. 36 (p. 201); hence in fig. 41, and similarly in fig. 42 below, the lateral principal maxima should have been drawn considerably lower than the central one. The effects of the diffraction bands arising from each slit by itself have not been shown either, as otherwise the figure would become too complicated.

Diffraction by a Large Number of Slits. — If we imagine the number of slits increased first to 8, then to 16, and finally more and

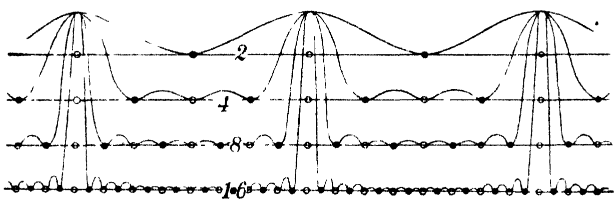


Fig. 42.—Variation of the maxima of intensity with the number of slits

more, the total effect may be obtained by a repetition of the process for four slits. We shall content ourselves by considering the final results, extending the drawing of fig. 41 for 8 slits and 16 slits. We thus obtain fig. 42, from which we see that multiplication of the slits causes the maxima to rise higher and higher above the axis of abscissæ and the points lying between them to lie closer and closer to the axis of abscissæ. The more the slits are multiplied, the more markedly does this occur, so that if the number of slits is very large we obtain sharply bounded, very narrow, and very bright interference lines on a dark background. The positions of these lines coincide with the positions of the maxima for two adjacent slits.

If two neighbouring slits are a distance b *apart (measured from the centre of one slit to the centre of the next), bright interference lines are produced at places where*

$$\sin\alpha = n\frac{\lambda}{b}; \quad n = 0, 1, 2, \ldots .$$

The central interference line, for which $\sin\alpha = 0$, lies at the position

O, where the convergent lens denoted by K in fig. 32, p. 199, would form an image of the source of light L, if the diffracting body GG were not present. The lines on either side are called *diffraction bands of the first, second, third, . . . , order.*

9. Diffraction Gratings.

A large number of slits very close together at regular intervals in an opaque screen forms what is called a **diffraction grating,** and the distance between the centres of successive slits, i.e. the quantity b in our above discussion, is called the **grating constant.**

A diffraction grating is produced by ruling a number of fine parallel *furrows* on a piece of plate glass or a flat reflecting metal plate by means of the diamond of a dividing machine. The light transmitted by the glass plate or reflected by the metal plate is then diffracted at these furrows. The *ridges* lying between the ruled portions of the glass plate let through the light like narrow slits; at the ridges of the metal plate there is regular reflection of light, as if the light from a source behind the mirror passed through by narrow slits.

With a good ruling engine it is possible to produce very accurate diffraction gratings. Excellent gratings were made on speculum metal* by Rowland in Baltimore (1882), with up to 20,000 lines to the inch, so that the distance between two successive lines, the grating constant b, is 0·00127 mm. When such a finely ruled grating is used, the angular separation of the interference bands is very great, for as b is very small $\sin\alpha = \lambda/b$ attains a high value for a definite wave-length λ.

If light from a slit, made parallel by a lens as in fig. 32, p. 199, is allowed to pass through a grating on glass or to be reflected from a reflection grating (the lines of the grating being parallel to the slit) and the rays from the grating are then made to converge on a screen by means of another lens, we obtain an image of the slit surrounded by the diffraction phenomena described in the previous section. From the equation

$$\sin\alpha = \frac{n\lambda}{b}$$

it follows that for small deviations the distances of the bright diffraction bands from the centre are approximately proportional to the wave-length.

In fig. 43 the position of the first five diffraction bands is shown for blue light of wave-length $\lambda_b = 400\ m\mu$ (on the left-hand side) and red light of wave-length $\lambda_r = 700\ m\mu$ (on the right-hand side), b, the constant of the grating used, being 0·005 mm. The positions of the diffraction bands are shown separately

* Very good gratings ruled on glass were produced by Nobert (Barth, Pomerania) reaching up to 400 lines to the millimetre. Rowland constructed three ruling engines; the first automatically ruled up to 1700 lines per millimetre, the second, which had fewer defects, ruled 20,000 lines to the inch, the third, 16,000 or a definite fraction of this number to the inch. More recently this number has been greatly exceeded, both on glass and on speculum metal, in gratings ruled by Lyle at Melbourne.

below the screen S. Generally speaking, the diffraction maxima are feebler the greater the order. Thus better separation is obtained at the cost of loss of intensity. By giving the rulings a special shape it is possible under certain circumstances to get a high percentage of the intensity of the diffracted light concentrated in one of the high orders, usually on one side only (WOOD's "echelette" grating).

When white light is used, every wave-length must give rise to a series of bands, the separate members of which lie between the blue and red bands marked by the same numeral. It follows that the complete diffraction band of the first order consists of a succession of all the colours from blue to red, i.e. a complete spectrum. This is indicated by the parallelogram in the bottom row of fig. 43. In actual fact the spectrum has the form of a rectangle, but is shown here as a parallelogram in order that the relative positions of the various orders of diffraction spectra may remain clear.

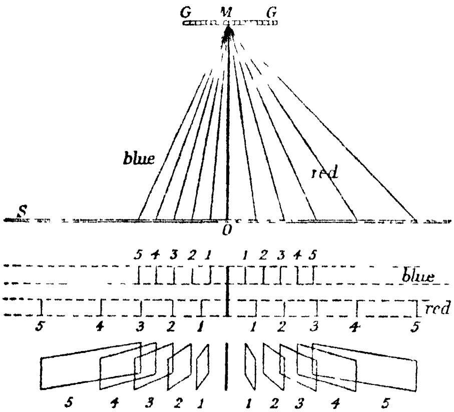


Fig. 43.—Fraunhofer grating spectra of various orders

The first spectrum (1) on either side is completely separated from the central maximum (of zero order) and also from the remaining spectra. The red end of the second-order spectrum (2) partly overlaps the blue end of the third-order spectrum. The red end of the third-order spectrum actually overlaps the spectrum of the fifth order. In the spectrum of the fourth order, therefore, there are no pure spectral colours at all, as it is overlapped partly by the third-order spectrum and partly by the fifth-order spectrum. With the spectra of higher order this overlapping occurs to a more pronounced degree; hence at a considerable distance from the centre the screen appears white, as it is illuminated by light of every wave-length.

We immediately recognize from the figure that the spectrum is longer the higher its order. Hence the various colours of the spectrum are separated the more (the resolving power is greater, that is, the distance between two lines with definite wave-lengths is greater) the higher the order of the spectrum.

If we confine ourselves to the first-order spectrum, we may equate sine and tangent; that is, the wave-lengths of the colours produced on the plane screen S are proportional to their distances from the diffraction band of zero order. If we know the wave-length of any colour and attach a proportionate scale to the screen so that its zero

lies at O, we can immediately read off all the other wave-lengths from it. A spectrum of this kind is called a **normal spectrum.**

Diffraction spectra are normal spectra.

If we use sunlight to produce a diffraction spectrum, Fraunhofer's lines * of course appear in it. On the coloured plate (facing p. 158) two spectra are shown one above another, the upper being a *dispersion spectrum* produced by a flint-glass prism and the lower a normal *diffraction spectrum*. In the dispersion spectrum the blue is much more drawn out than in the normal spectrum, while the red is much more crowded together. On the other hand (see fig. 14, p. 160), if we do not use the wave-length as the co-ordinate, but the frequency ν ($= c/\lambda$), which is more important theoretically, the relationship departs less from linearity with the prismatic spectrum than it does with the diffraction spectrum.

The Grating Spectroscope (fig. 44).—This apparatus consists of a tube carrying a slit S which is illuminated by light from the source under investigation. The rays are then made parallel by means of the collimating lens B and fall on the glass grating GG. The position of the diffraction band of zero order is indicated by the arrow at O. The angle at which the parallel rays forming the first-order band at A in the focal plane of the telescope emerge is given by the equation

$$\sin \alpha = \frac{\lambda}{b}.$$

The diffraction band is observed through C, the eyepiece of the telescope. A wave-length scale may be directly attached to the telescope stop at A and the wave-length read off. If we wish to photograph the spectrum, we have merely to replace the telescope by a camera focused for infinity.

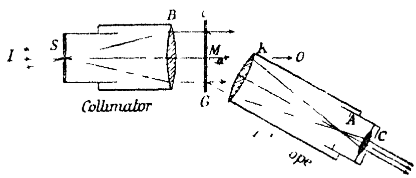


Fig. 44.—Spectroscope and grating

In their passage through glass the rays are partly absorbed, particularly the ultra-violet rays (Chap. VII, § 4, p. 160). Hence in investigations of this part of the spectrum it is impossible to use either glass lenses or a glass grating. For such investigations Rowland's *concave* reflection gratings are particularly suitable, as they produce a sharp image of the spectrum directly by reflection.

These concave gratings require to be set up in a particular way. In fig. 45 GG denotes the concave grating and A its centre of curvature; A is accordingly the image of itself in the concave mirror. A circle with centre C is drawn on AM as diameter. A slit S, which is illuminated by the source (L) under investigation, lies on this circle. The rays of light from the slit fall on the concave

* Their wave-lengths were first measured by FRAUNHOFER in 1823 with the help of gratings.

mirror and produce an image of the slit at O. The point O is accordingly illuminated as if a source at L′ were to send its light through the grating GG; hence O may be regarded as a diffraction grating of zero order. The position of the first-order spectrum A is then such that

$$\sin \text{OMA} = \sin\alpha = \frac{\lambda}{b}.$$

Now ∠AOM is a right angle, so that

$$\sin \text{OMA} = \frac{\text{AO}}{\text{AM}} = \frac{\text{AO}}{r},$$

where r is the radius of curvature of the concave grating. From the two equations it follows that

$$\frac{\text{AO}}{r} = \frac{\lambda}{b}, \text{ i.e. AO} = \lambda\frac{r}{b}.$$

AO is therefore proportional to the wave-length λ. If we replace AO by the arc of the circle, we find that a wave-length scale may be directly fixed to the circle, i.e. that a normal diffraction spectrum is formed at A. A curved photographic plate P is then placed at A and directly receives the normal spectrum of the source L.

Note.—The above argument includes a number of simplifications and omissions, but the same result, namely, that a sharp normal diffraction spectrum is produced at A when the apparatus is set up as shown in the figure, is obtained when the calculation is carried out on stricter lines.

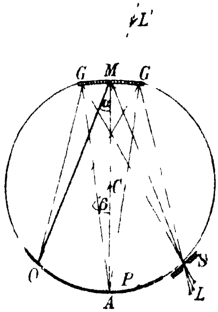


Fig. 45.—Adjustment of a Rowland concave grating

The Importance of the Measurements obtained with the Diffraction Grating.—The extreme importance of the diffraction grating lies mainly in the fact that from the determination of the grating constant, the dimensions of the apparatus, and the spectra obtained, i.e. from measurements of length which can be carried out very accurately, it is possible to measure the wave-lengths of the lines occurring in spectra directly with great accuracy. Prismatic spectrometers can then be calibrated on the basis of the values so obtained. There is also the further advantage that by using concave reflecting gratings it is possible to avoid the use of lenses, and hence to let the rays travel entirely in a vacuum; by this means it is possible to penetrate into regions of the spectrum which, owing to absorption, would otherwise be inaccessible. Actually almost the whole region of the spectrum of electromagnetic waves given in fig. 47 (p. 646) of Vol. III has been investigated and wave-length data have chiefly been collected in this way. For accurate measurements in the visible region and adjacent regions, interferometric methods (p. 193) are steadily gaining in importance.

From the fundamental formula governing the action of the dif-

fraction grating we see that the grating constant must be of a magnitude suited to the wave-length. It must not be too small, as otherwise the above interference phenomena do not occur at all, and it must not be too large, as otherwise it is not possible to separate the spectra from the central image.

By letting the waves fall on the grating almost at grazing incidence, however, it is possible to get a good separation of the various orders of spectra by reflection, even when the constant of the grating is relatively large. For example, by using the total reflection of X-rays (fig. 19, p. 165) we can in this way investigate X-ray spectra (fig. 46, Plate XVI) with ordinary glass gratings (by reflection). The very exact measurements of wave-length which can be made in this way are of particular importance in that they enable AVOGADRO's number and the charge on an electron (see Vol. V) to be determined by a method depending solely on measurements of length and density.

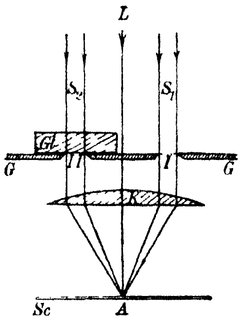


Fig. 47.—To illustrate the echelon grating

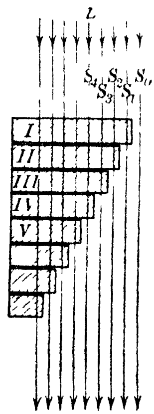


Fig. 48.—Michelson's pile of plates

Michelson's * Echelon Grating.—It follows from the discussion based on fig. 43, p. 206, that the diffraction spectra of higher order are more spread out than those of lower order, i.e. the resolving power is greater in the former case (p. 211). Diffraction spectra of high order, however, are very faint; moreover, if gratings with a large number of lines and a small grating constant are used, they are very far to one side or the other. In a Rowland grating with very close ruling the third-order spectrum does not occur at all, for the expression $\sin\alpha = n\lambda/b$ then takes a value for $n = 3$ which is greater than unity. These two difficulties prevent the use of the higher-order spectra

* See p. 190.

produced by diffraction gratings. Starting from the idea that the diffraction spectra of high order are produced by pencils of rays with a large difference of path, MICHELSON developed his *echelon grating*.

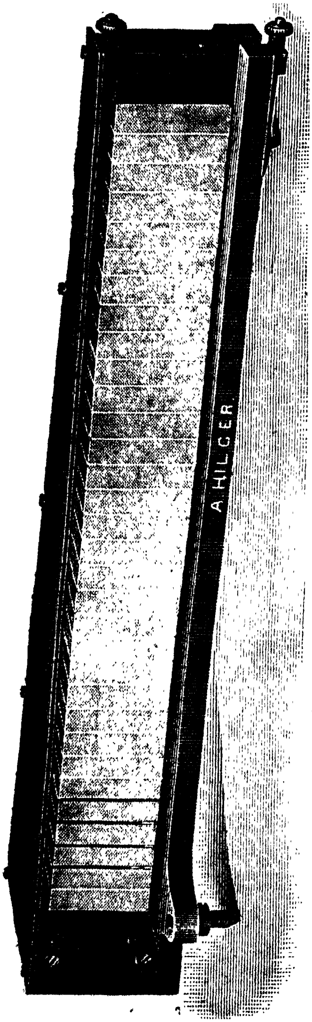
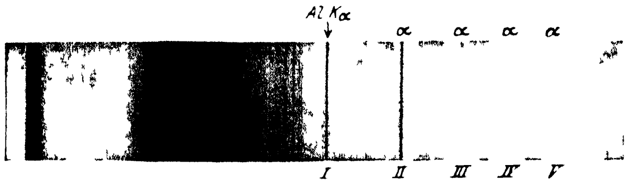


Fig. 49.—An echelon grating (Adam Hilger, Ltd.)

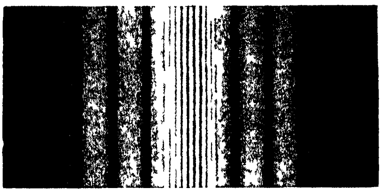
To understand its action we begin by considering fig. 47. Here I and II denote two slits in the opaque screen GG. If two pencils of parallel rays, S_1 and S_2, from the source L at infinity pass through the slits I and II, they are made to converge to the band of zero order A on the screen Sc by the condensing lens K. If, however, we place the flat glass plate Gl of thickness d over slit II, the pencil S_2 is retarded in phase by the amount $\Delta w = d(\mu - 1)$, where μ is the refractive index of the glass. Two pencils of parallel rays with the difference of path Δw accordingly interfere at A and produce an image which exactly corresponds to a diffraction band of high order.

MICHELSON made use of this property in his echelon grating, shown diagrammatically in fig. 48. This consists of several glass plates I, II, III, . . . one on top of another; these are as nearly as possible of the same thickness, and each overlaps the preceding by the same small amount. As a result each pencil of parallel rays starting from L is subject to a retardation corresponding to the optical thickness of the plate of glass traversed by it. Thus the pencils emerge with constant differences of path between adjacent pairs. When they are combined by means of a convergent lens a bright interference pattern is produced, which corresponds to a diffraction spectrum of very high order—say the ten-thousandth. The external appearance of an echelon grating is shown in fig. 49. Fig. 50 (Plate XVI) shows a photograph of the green mercury line (see the coloured plate) taken with an echelon grating (cf. fig. 30*b*, Plate XV).

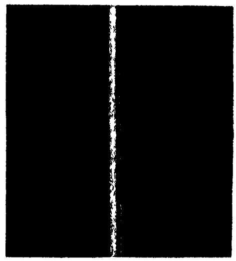
The Resolving Power of Spectroscopes.—If light consisting of two



Ch VIII Fig 46 Photograph of an X ray line from aluminium (Kα λ 8 333 Å) by Backlin taken with a glass grating and showing five orders The feebler lines to the left of the first order aluminium line are due to the oxide cathode (From *Handbuch der Experimentalphysik* Vol 24 Part II (Akademische Verlagsgesellschaft Leipzig))



Ch VIII Fig 38 Diagram of the interference pattern formed by a double slit (p 201)



Ch VIII Fig 50 The green mercury line λ 5461 Å resolved by the echelon grating (1 mm ~ 0 05 Å)

Ch X, Fig 36 —Double refraction due to strains in rapidly cooled glass (p 247)

Ch X Fig 37 —Artificial double refraction used to investigate distributions of stress (a piece of glass under compression supported at one point on the right and at a number of points on the left)

colours which differ very little in wave-length is investigated in a spectroscope, the two are only recognized as different kinds of rays if their difference of wave-length $d\lambda$ reaches a certain magnitude. If the mean wave-length of the rays is λ, $\lambda/d\lambda$ is called the *resolving power* of the spectroscope. A resolving power of 100,000 accordingly means that rays which differ by 1/100,000 of their wave-length can still be recognized as separate.

Prismatic Spectra.—Two adjacent spectral lines (monochromatic images of the slit) cannot be separated to an indefinite extent by introducing greater magnification in observing or photographing the spectrum. For owing to the diffracting effects of the slit each spectral line has a finite breadth, which by pp. 16, 145 is smaller the larger the slit. Hence to increase the resolving power two possibilities are available: (1) using large lenses and prisms and (2) increasing the dispersion (usually by employing a larger number of prisms, but not as a rule more than three). The first method is not advisable once a certain degree of resolving power has been attained, owing to the various aberrations, which are usually increased likewise.

Taking diffraction and dispersion into account, we obtain (according to LORD RAYLEIGH)

$$\frac{\lambda}{d\lambda} = C\frac{d\mu}{d\lambda},$$

where C is the length of the base of the prism and $d\mu/d\lambda$ the dispersion.

Example.—With a flint-glass prism ($\mu = 1{\cdot}65$) having a side of 1 cm. the two components of the D line ($d\lambda = 6$ Å.) can just be separated.

Interference Spectra.—The interference maxima become narrower the greater the number (p) of separate interfering elements (lines on a grating, plates in an echelon grating, number of reflections in the Lummer-Gehrcke plate). The distance between the maxima corresponding to two rays of different wave-length is greater the higher the order (m) of the interference phenomenon observed. It follows that

$$\frac{\lambda}{d\lambda} \sim mp.$$

In ruled gratings p is made as great as possible, while m cannot as a rule exceed 2 or at most 3. In the other types of apparatus m is the chief effective factor. Table III (see p. 212) gives a comparison of the resolving powers attained hitherto with various kinds of instruments for a wave-length of 500 $m\mu$. As the wave-length increases the resolving power of prism spectroscopes gets worse, whereas that of interferometers improves, and conversely for smaller wave-lengths.

TABLE III

	Resolving Power	Order of Spectrum
Flint prism (C = 10 cm.)	10,000	—
Ruled grating (p = 100,000: 3rd order spectrum)	300,000	3
Echelon grating (40 plates, each 1 cm. thick) ..	400,000	10,000
Lummer-Gehrcke plate (1 cm. thick, 20 cm. long)	800,000	40,000
Air plate (20 cm. thick)	7,000,000	800,000

CHAPTER IX

The Velocity of Light

1. Methods for Determining the Velocity of Light.

The Eclipses of Jupiter's Moons.—The velocity of light was first determined by the Danish astronomer OLE RÖMER * in 1675. He calculated it from the **retardation of the onset of the eclipses of a satellite of Jupiter** during the period in which the earth in its motion round the sun is receding from Jupiter.

Let the circle with centre S in fig. 1 represent the earth's orbit and the small circle with centre J the orbit of the first principal satellite of Jupiter, which has an orbital period of 1·769 days or about $42\frac{1}{2}$ hours. When the earth is at I or at III the orbital period (i.e. the time between two successive eclipses of the satellite) as observed by a terrestrial observer agrees exactly with the true orbital period of the satellite. The observer, it is true, does not see the onset of an eclipse at the instant at which it really takes place, but later, the interval representing the time taken by light in travelling from Jupiter to the earth; but the retardation is the same for two successive eclipses, so that the interval between them agrees with the actual orbital period. If, on the other hand, the earth is at II, i.e. is receding from Jupiter, the orbital period of the satellite appears longer; for while the satellite has gone once round Jupiter the earth has moved a considerable distance away from Jupiter. Two successive eclipses are again observed later than they actually occurred, but as the light in the case of the second eclipse has farther to travel than in the case of the first, the observation is more belated in the second case than in the first. Hence the orbital period is apparently lengthened. The same occurs at every revolution of the satellite during the time that the earth is moving from I to III by way of II. On the other hand, during the time that the earth is moving from III to I by way of IV each orbital period is shortened. These separate increases in the orbital period on one side of the earth's orbit (I–II–III) and decreases on the other side of the earth's orbit (III–IV–I) are additive. If now, starting from the instant

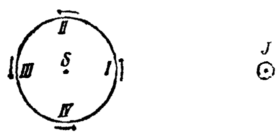


Fig. 1.—Various positions of the Earth, the Sun, and Jupiter

* OLE RÖMER (1644–1710), a Dane, was assistant to DOMENICO CASSINI, the director of the Paris observatory, from 1669 to 1675. During this time both of them continued the tables of the eclipses of Jupiter's moons begun by CASSINI in 1660. The fact that the eclipses are retarded, and his explanation of the fact, were communicated to the Paris Academy by RÖMER on 22nd November, 1675; but the members of the academy, like CASSINI, were not disposed to accept the explanation of the phenomenon as due to light being propagated with a finite velocity.

I, we calculate from the true orbital period the instant at which an eclipse should occur when the earth is in the neighbourhood of III, it is found on observation that the beginning of the eclipse is 1000 seconds late. If, starting from the instant III, we calculate from the true (i.e. the mean observed) orbital period the instant at which an eclipse should occur when the earth is at I, it is found on observation that the beginning of the eclipse is 1000 seconds too early. This must be the time taken by the light to travel from I to III, i.e. right across the earth's orbit. The diameter of the earth's orbit is 300,000,000 Km.; hence in free space light traverses one-thousandth of this distance in one second, i.e. 300,000 Km.

Fizeau's Method.—The velocity of light was also determined by J. Bradley in 1725 from the so-called aberration of the fixed stars. His method will be described in Vol. V. Later Fizeau * measured the velocity of light by means of *terrestrial* experiments (1849). At the ends of a base-line 8633 m. long two telescopes I and II (fig. 2)

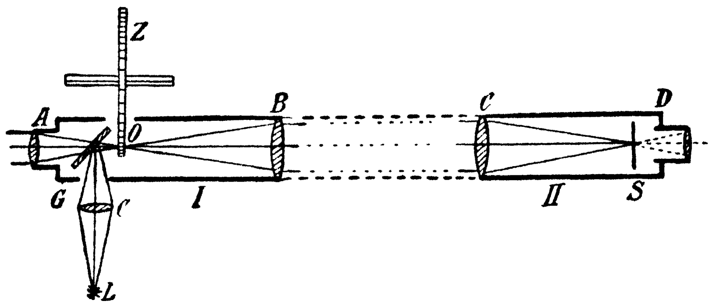


Fig. 2.—Fizeau's apparatus for measuring the velocity of light

focused for infinity were set up facing one another in such a way that an observer looking through one telescope saw the cross-wires of his own telescope covering the cross-wires of the other telescope. The cross-wires of the observer's telescope I were then replaced by the rim O of a toothed wheel Z (the teeth and the gaps being of the same width) which could be made to revolve rapidly about its axis, which was parallel to that of the telescopes. Between A, the eye-piece of the observer's telescope, and O, the edge of the toothed wheel, was inserted a plane unsilvered piece of plate glass G, inclined at an angle of 45° to the axis of the telescope. The point source of light L outside the telescope emitted rays which were made to converge by a lens C and were reflected from the glass plate in such a way that a real image of L was formed in the plane of the toothed wheel. If the wheel were *at rest* and the light fell on a gap between the teeth, the rays continued to diverge and, after being made parallel by the objective B, left the telescope I in the direction of the axis of both

* H. Fizeau (1819–96), Professor of Physics in Paris, made a number of noteworthy investigations in optics. He also suggested the use of a condenser in the induction coil (Vol. III, p. 517).

telescopes. The pencil of rays then entered the objective of telescope II, was made to converge by it, and fell on a plane metal mirror S which had been substituted for the cross-wires of telescope II. After reflection the rays of the convergent pencil travelled over their former path in the reverse direction and could then be observed through A, the eyepiece of the telescope I, and the transparent glass plate G.

If the toothed wheel Z were then set rotating the field of view was found to become dark when the rate of revolution reached a certain value. For if the toothed wheel had turned through an amount exactly corresponding to the breadth of a tooth during the passage of the pencil of rays backwards and forwards, a pencil of rays which passed through a gap on its outward journey would fall on the next tooth on the return journey. When the rate of revolution was doubled the field of view was found to become bright again, since the light which passed through a gap on its outward journey could then pass through the next gap on the return journey. As the rate of revolution was increased still further the field of view became alternately bright and dark.

Fizeau used a wheel with 720 teeth. Darkness first appeared when the wheel rotated 12·6 times per second. Then the time necessary for a gap to be replaced by the tooth following it was $1/(2 . 720 . 12{\cdot}6) = 1/18{,}150$ sec. During this time the light must have traversed the distance between one telescope and the other twice, i.e. a distance of 2×8633 m. or 17·27 Km. That is, in one second light would traverse a distance of $18150 \times 17{\cdot}27$ Km. $= 313{,}000$ Km.

Perrotin in 1901, using the same method with improved instruments and a base-line of 46 Km., obtained the value (corrected to free space *) $299{,}860 \pm 80$ Km.

Foucault's Method.—A few years after Fizeau had determined the velocity of light experimentally for the first time by the method just described, Foucault † published a second method (1862), which enabled him to measure the velocity of light by means of a rotating mirror. The most characteristic part of his apparatus is a small plane mirror which may be set in rapid rotation about a vertical axis. By means of a small turbine actuated by compressed air Foucault obtained rates of over 800 revolutions per second. He determined the number of revolutions per second from the sound produced.

The arrangement of Foucault's apparatus is shown diagrammatically in figs. 3 and 4. S is the mirror, which rotates about the axis O. Sp is a slit illuminated by a strong source of light L in conjunction with the condensing lens C, or by sunlight. K is a convergent lens which if the mirror S were not there would

* If c_A is the velocity in air, c that in free space, and $\mu = 1{\cdot}00028$ the absolute refractive index of air (p. 54), $c = c_A \times 1{\cdot}00028$.

† Léon Foucault (1819–68), originally a physician, was subsequently a member of the Paris Academy.

produce a real image of the slit at B_1. The mirror S, which at first we shall assume to be at rest, reflects the light rays so as to produce the real image of the slit at the point B on the concave spherical mirror H. The centre of curvature of the concave mirror H is to lie on the axis O about which the mirror S rotates;

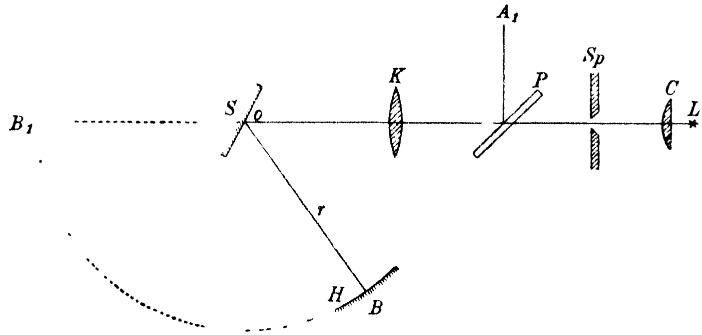


Fig. 3.—Foucault's apparatus

it will then reflect all the rays falling on it from S back on to S. The returning rays accordingly give rise to a real image of the slit, which coincides with the slit itself. The central ray of the pencil forming the image is given by SpOBOL

We have still to take into account the action of the plane unsilvered glass plate P inclined at 45° to the path of the rays. The plate reflects part of the light in a direction at right angles to its former path. These reflected rays give rise

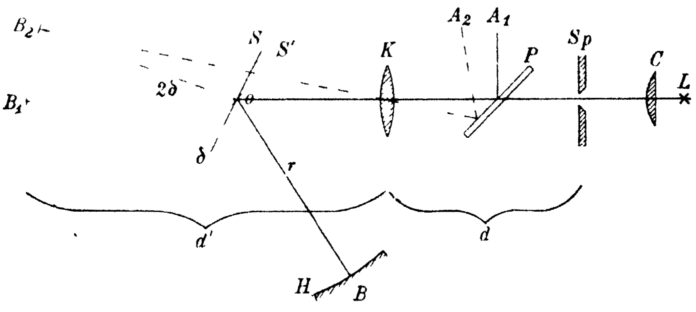


Fig. 4.—Measurement of the velocity of light by means of a rotating mirror

to a real image of the slit at A_1, which is observed by means of a microscope with an eyepiece micrometer.

If the mirror S is set revolving slowly, the image A_1 only appears when the rays reflected by S actually fall on the concave mirror. It follows that the image of the slit will periodically appear and disappear. If we gradually increase the rate of revolution, the image will flicker at first, but appears steady when the number of revolutions per second reaches even the low value of 10.

The rate of revolution of the mirror is now increased so greatly (fig. 4) that during the time required by light to travel from S to the concave mirror H and back again to S the mirror S has revolved through a small but finite angle δ

so as to occupy the position S′. As a result the virtual image of the point B formed by the mirror S′ will lie not at B_1, but at B_2 (p. 34, fig. 4). The prolongation of B_2O is the reflected ray corresponding to BO after the mirror has rotated, and $\angle B_1OB_2 = 2\delta$. The rays reflected by S′ then appear to come from B_2; they pass through the condensing lens K and now produce a real image of the slit at A_2. This image of the slit is displaced relative to the image A_1 through an amount e which may be read off on the eyepiece micrometer of the observer's microscope.

If we denote the radius of curvature of the concave mirror by r, the distance of the slit Sp from the condensing lens K by d, the distance of the image B_1 from the condensing lens by d', the number of revolutions of the mirror S per second by n and the velocity of light by c, we have the following relationships:

The angular velocity of the rotating mirror is $2\pi n$. As the time taken by the light to traverse the path OB and back, i.e. a distance $2r$, is $2r/c$, the angle δ through which the mirror rotates during this time is given by

$$\delta = 2\pi n \cdot \frac{2r}{c} = \frac{4\pi n r}{c};$$

hence

$$\angle B_1OB_2 = 2\delta = \frac{8\pi n r}{c}.$$

As the angles are very small the ratio of the angle B_1KB_2 and the angle B_1OB_2 is the reciprocal of the ratio of the corresponding lines r and d'; hence

$$\angle B_1KB_2 = \frac{8\pi n r}{c} \cdot \frac{r}{d'} = \frac{8\pi n r^2}{c d'}.$$

As the slit, and hence also the images of the slit A_1 and A_2, are at a distance d from the condensing lens K, measured along the light rays, it finally follows that the interval $A_1A_2 = e$, through which the image of the slit observed in the microscope is displaced, is

$$e = d \cdot \angle B_1KB_2 = \frac{8\pi n r^2 d}{c d'}.$$

From this expression, in which all the quantities but c are directly measurable, it follows that

$$c = \frac{8\pi n r^2 d}{e d'}.$$

Example.—A concave mirror with a radius of curvature of 9 m. (9×10^2 cm.) was used, and a condensing lens was set up so that d was 2 m. (2×10^2 cm.) and d' 10 m. (10×10^2 cm.). When the mirror reached 800 revolutions per second, the displacement of the image of the slit was 1·08 mm. ($1{\cdot}08 \times 10^{-1}$ cm.).

Substituting these values in the formula above, we have

$$c = \frac{8 \times 3{\cdot}14 \times 8 \times 10^2 \times 81 \times 10^4 \times 2 \times 10^2}{1{\cdot}08 \times 10^{-1} \times 10 \times 10^2} = 3 \times 10^{10} \text{ cm./sec.}$$

A. A. Michelson's Method.—The mirror method has been brought to an extraordinary degree of perfection by MICHELSON (1926). By skilful improvements in the apparatus he has been able to increase the brightness and steadiness of the image to such an extent that it is possible to make the light traverse a distance of about 70 Km.

The lay-out of the apparatus is shown in fig. 5. Light from a very bright arc lamp (such as the Beck lamp mentioned on p. 25) passes through the slit S and falls on an eight-sided prism of glass or nickel steel. Its surfaces are made as highly reflecting as possible and are so ground that the angles between the sides of the prism are identical to within a millionth part of their magnitude. This mirror-prism is capable of rotating about an axis at right angles to the paper, and is driven by compressed air. To begin with, we shall assume that the prism is at rest. The pencil of rays incident at *a* is reflected to the fixed plane mirror *b* and thence to an obliquely-placed plane mirror *c*, which reflects the rays on to a concave mirror H_1 of diameter 60 cm. in such a way that they travel

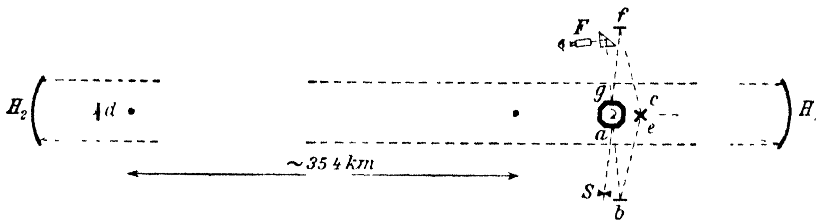


Fig. 5.—Michelson's apparatus for measuring the velocity of light

as parallel rays to the exactly similar concave mirror H_2 at a distance of about 35·4 Km. There they are thrown on to a small plane mirror *d* at the focus of the concave mirror, are reflected back to the concave mirror, travel back to H_1, and fall on a small plane mirror *e* placed just below *c* but almost at right angles to it. Thence the light is reflected by way of the plane mirror *f* to the side *g* of the "octagon", and then by way of a totally reflecting prism to the observer's telescope F, which is provided with an eyepiece micrometer. When the rotating mirror is at rest the image of the slit S is accordingly seen in the telescope F. If the mirror is set rotating, the image vanishes, as the light reflected from *a* at the instant at which it occupies the position shown in the figure, on arriving at *g*, finds this mirror already turned through a definite angle, owing to the time which the light requires to travel over the distance of 70 Km. If the mirror is now rotated more rapidly, there comes an instant when the image of the slit again becomes visible, i.e. when at the instant of arrival of the flash of light the next face of the prism exactly occupies the position of *g*, i.e. the prism has rotated through exactly one-eighth of a complete revolution. If the path over which the light travels is 70·8 Km. this occurs at about 528 revolutions per second. The experiment now consists in keeping the mirror rotating at exactly this value, which it is possible to do for several seconds by comparison with an electrically-driven tuning-fork. This time is sufficient to measure the small remaining displacement of the slit relative to the position of rest, by means of the eyepiece micrometer of F. The base-line was measured by a staff of expert surveyors with an accuracy of 1 in 5 millions. The tuning-fork was calibrated with equal accuracy by means of a standard clock.

The value obtained for the velocity of light in a vacuum is

299,796 ± 4 Km./sec.

More recent experiments, in which the path traversed by the light is over 260 Km., have not yet been concluded.

For the determination of the velocity of light by means of the

Kerr cell see p. 258. The velocity of long electric waves, and on the other hand that of X-rays, has also been found to be 3×10^{10} cm./sec. in a vacuum, within the limits of experimental error.

2. The Velocity of Light in Material Media.

Measurement of the Velocity of Light in Water.—FOUCAULT inserted a tube full of water in the path of the rays between S and H (fig. 3). This enabled him to measure the velocity of light in water also. FOUCAULT's experiments yielded the result that the *velocity of light in water is only three-fourths of the velocity of light in air*. The ratio of the velocities of light is exactly equal to the ratio of the refractive indices.

FOUCAULT's experiments are of particular importance because they decided the question whether the velocity of light in optically denser media is less than or greater than the velocity in optically less dense media. This question is closely connected with the question whether light consists, as on NEWTON's emission theory,* of small particles (corpuscles), which are thrown out in all directions by the source of light and proceed in straight lines, or whether light is propagated like a wave motion. FOUCAULT's experiments decided the matter finally in favour of the wave theory of light. Here, however, it is to be noted that recent results on the properties of light quanta show that it is possible to reconcile the emission theory with FOUCAULT's experiments (Vol. V).

Group Velocity and Wave Velocity.—The methods described above for determining the velocity of light do not really measure the same quantity in the same way. For the methods of FIZEAU and ROMER apparently give the velocity with which a sharply bounded train of waves, a *wave group*, advances. FOUCAULT's method, as closer investigation shows, also gives the *group velocity* (p. 172); BRADLEY's aberration method is the only one that gives the *wave velocity*, according to the usual conception of the propagation of light. Now in a vacuum light of every wave-length has the same velocity; there is no dispersion. Measurements by SHAPLEY in 1923 showed that the variations of intensity of stars which are so distant that their light takes 40,000 years to reach the earth take place simultaneously for green light and for violet light, so that within the limits of experimental error the *velocity of light* is the *same* for these two colours throughout the *universe* (in free space) to within one part in 2×10^{10}. Hence in free space the group velocity and the wave velocity have the same value (Vol. II, p. 226). This is not so, however, for media in which there is *dispersion* (p. 171). Carbon disulphide, for example, is a medium with a high dispersion. In it the wave velocity and the group velocity differ markedly. In fact, A. MICHELSON (1884) was greatly

* DESCARTES in his *Dioptric* (1637), extending PLATO's ideas, had previously developed the theory that light is to be regarded as consisting of moving corpuscles, but not in a very clear way. Using this idea, he succeeded in deducing the law of refraction. The conclusions which DESCARTES drew from his version of the law of refraction were immediately attacked by FERMAT (p. 221), who was able to suggest another basis for the law of refraction and obtained the more intelligible hypothesis that light should be propagated more slowly in the optically denser medium.

surprised to obtain the value $c'' = c/1{\cdot}77$ for the velocity of light in carbon disulphide, using Foucault's method, the value calculated from the refractive index being only $c' = c/1{\cdot}64$, where c is the velocity of light in air. Now by Vol. II, p. 226, the group velocity c'' is given by $c' - \lambda\, dc'/d\lambda$, where c' is the wave velocity, varying with wave-length, in the dispersive medium. For carbon disulphide measurements of the dispersion give $\frac{\lambda}{c}\frac{dc'}{d\lambda}$ equal to 0·075. Hence for the group velocity we have

$$c'' = c' - \lambda \frac{dc'}{d\lambda} = c' - c' \times 0{\cdot}075 = c'(1 - 0{\cdot}075).$$

Inserting $c' = c/1{\cdot}64$, we have

$$c'' = \frac{c(1 - 0{\cdot}075)}{1{\cdot}64} = \frac{c}{1{\cdot}77}$$

in complete agreement with Michelson's result. By taking dispersion into account it is accordingly possible to explain the apparent inconsistency of the experimental data, which had been further confirmed by very accurate experiments, especially by GUTTON (1911).

Measurements of Wave-length.—The variation in the velocity of propagation of light with the refractive index of the medium may also be demonstrated by direct measurements of wave-length. If the film of air between the lens and the glass plate in the experiment of Newton's rings is replaced by water, by letting a drop of water flow in between the lens and the plate, and all the other experimental conditions remain the same, the rings move closer together, the radius of each ring shrinking to 13/15 of its original value. As the frequency of the light used remains the same, this can only be due to a change in wave-length. Hence the wave-length in water is only about three-quarters of the wave-length of the same light in air; for the wave-length is proportional to the thickness of the film, i.e. to the square of the radius of the ring, and $(13/15)^2 = 169/225 \approx 3/4$.

As by the equation $c = \nu\lambda$ the velocity of propagation is proportional to the wave-length, it follows that the velocity of light in water is only three-quarters of the velocity of light in air; it is accordingly about 225,000 Km./sec.

The same result is obtained from measurements with a diffraction grating.

A rectangular metal box K without a lid (fig. 6), 1 m. long and bounded at its smaller faces by flat pieces of plate glass, is set up so that the centre of the box is at the same height as the condensing lens C of a mercury lamp B, whose light passes through a green filter and is thus rendered monochromatic. Immediately in front of one end of the box (the right-hand end in the figure) there is a diffraction grating G, which has the constant $d = 0{\cdot}05$ mm. The illuminating slit and the collimating lens L are placed between the box and the condensing lens, while the ground-glass screen S is fixed in the box so that diffraction bands

are distinctly formed on the screen. If the box is then half filled with water, two systems of diffraction bands are formed on the screen S. The upper system of bands is produced by light waves which have travelled through air, the lower

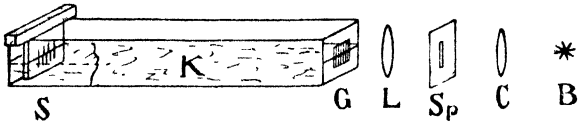


Fig. 6 —Grimsehl's apparatus for demonstrating the different velocities of light in different media

by light waves which have travelled through water. The bands of the two systems lie one on top of the other in such a way that four intervals between bands for water correspond to three intervals for air.

The distances between the diffraction bands are proportional to the wave-lengths and these to the velocities of light in the two media, so that the velocities of light in air and water are in the ratio of 4 to 3.

3. Fermat's Principle.

Fermat's * principle is usually stated as follows:

The path of a ray of light from point to point is always such that the time taken by the light to traverse it is a minimum.

Reflection. Two points A and B lie in front of the plane mirror SS′ (fig. 7). If these points are joined to two points C and F lying on the surface of the mirror, of which C is so situated that the ray ACB satisfies the law of reflection, but the ray AFB does not, it is required to prove that light takes less time to traverse the path ACB than it would do to traverse the path AFB.

If from F we drop the perpendicular FD on AC and the perpendicular FE on BC produced,

$$\angle BCL = \angle ACL$$

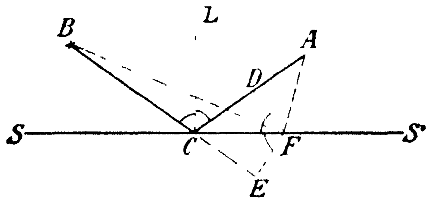


Fig. 7.—The law of reflection and Fermat's principle

by the law of reflection; hence the triangles CDF, CEF are congruent, so that

$$CD = CE.$$

Hence the path ACB is equal to AD + EC + CB or AD + EB. Further, AD < AF and BE < BF; hence

$$AD + EB < AF + FB,$$

that is, the path ACB is shorter than the path AFB. Hence we have Fermat's principle, that in reflection at a plane mirror the light traverses the path ACB in a shorter time than it traverses any other path AFB.†

* Pierre de Fermat (1608–65), a lawyer and member of the parliament of Toulouse, was at the same time one of the foremost mathematicians that France has produced; he brought forward the principle stated above in order to prove Snell's law of refraction.

† The fact that in reflection the path taken by the light is a minimum was actually known to Hero of Alexandria.

Refraction.—Let WW′ in fig. 8 represent the plane boundary between two media which differ in their refractive properties, e.g. air and glass. Let the point A be in air and the point B in glass. A and B are joined to two points C and F on the boundary surface such that ACB is the path of the light ray according to the law of refraction, while AFB is any other path. It is required to prove that light takes less time to traverse the path ACB than it would to traverse the path AFB.

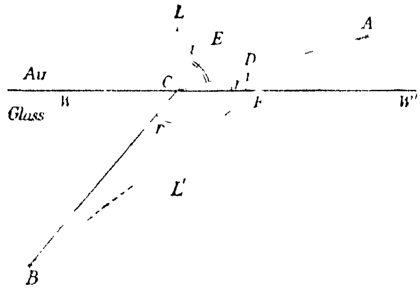


Fig. 8.—The law of refraction and Fermat's principle

If the refractive index of glass relative to air is μ, and if we take the velocity of light in air as unity, the velocity of light in glass is equal to $1/\mu$. To traverse the path ACB the light requires the time $AC \times 1 + CB \times \mu$; to traverse the path AFB it would require the time $AF \times 1 + FB \times \mu$. If i and r are the actual angles of incidence and refraction, $\sin i = \mu \sin r$.

If from F we drop the perpendicular FD on AC and the perpendicular FE on BC produced, we have

$$\angle DFC = i \text{ and } \angle CFE = r.$$

i.e. $\sin DFC = \mu \sin CFE$. As $\sin DFC = CD/CF$ and $\sin CFE = CE/CF$, it follows that $CD = \mu CE$.

The time which the light ray takes to traverse the path ACB corresponding to the law of refraction, i.e. the time to traverse $AC + \mu CB$, may be replaced by the time taken to traverse $AD + DC + \mu CB$.

Then as $DC = \mu CE$, it follows that the path $AC + \mu CB$ is equivalent to

$$AD + \mu CE + \mu CB = AD + \mu EB.$$

Now $AD < AF$, $EB < FB$, so that

$$AD + \mu EB < AF + \mu FB.$$

The last expression represents the time which the light would take to traverse the path AFB. Hence Fermat's principle is established for refraction also.

Note.—It is to be noted that under certain circumstances the path of the light may be a *maximum*. The *general* form of Fermat's principle is as follows:

Among all the neighbouring possible paths the path actually followed by the light is a maximum or a minimum.

CHAPTER X

The Polarization of Light

1. Polarization by Reflection.

The Fundamental Experiment.—The following experiment reveals an entirely new property of light.

A pencil of parallel rays (fig. 1) falls obliquely on a glass plate P which has its rear side blackened. The part of the light entering the glass is absorbed at the back of the plate, the part (M) reflected at the front surface alone remaining accessible to observation. If the reflected light M is allowed to fall on a second glass plate A which likewise has its rear side blackened, the light is again reflected in the direction N. If the second reflecting plate A is then rotated about the incident ray M as axis so as to occupy successively the positions shown in figs. 1–4, the intensity of the reflected ray varies four times. It reaches its maximum when the two planes of incidence, i.e. the planes containing the incident and reflected rays in each case, coincide (as in figs. 1 and 3) and its minimum when the two planes of incidence are at right angles to one another (as in figs. 2 and 4).

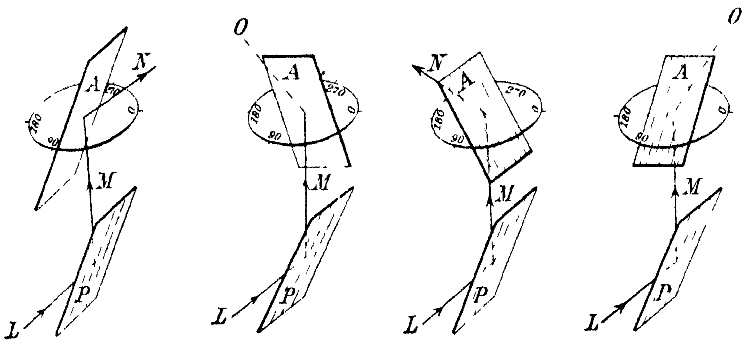


Fig. 1 Fig. 2 Fig. 3 Fig. 4

Polarization of light by reflection

If the light rays are allowed to fall on the two glass plates at *different* angles of incidence, the variation of the intensity of the light

reflected from the second plate is most marked when the angle of incidence is 55° for each mirror (fig. 5).

Reflection at the glass plate affects the pencil of light rays in such a way that it behaves quite differently in two directions at right angles to one another. This change in the properties of light is called **polarization**, the light so altered, **polarized light**, and the plane in which the light polarized by reflection is incident, the **plane of polarization**;* the apparatus by which polarization is brought about is called a **polarizer**.

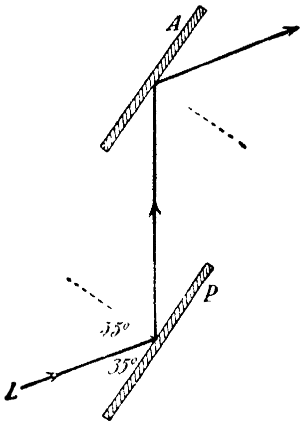


Fig. 5.—Light polarized by reflection. The plane of reflection is the plane of polarization

The Plane of Polarization and the Direction of Vibration.—The extreme importance of the discovery of polarization is due to the fact that the *transverse* nature of the motion which constitutes light was thereby established. In order to explain the behaviour of polarized light described above, we must assume that there is a difference between the two sides of the ray (if we imagine ourselves looking along the direction of the ray), that is, that the periodic variations of which light consists take place in a definite plane, which is determined by the direction of the ray and a direction at right angles to this (fig. 6). A vibration of this type taking place at right angles to the direction of propagation, that is, in which the variation resembles the motion of the particles of a vibrating string at right angles to the direction of propagation of the waves, is a *transverse* vibration.

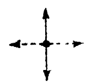


Fig. 6.—Difference of a polarized light ray in two planes

* The discovery of the polarization of light goes back to MALUS (1775–1812). One evening in 1808 he was looking through a piece of Iceland spar (p. 233) at the windows of the Luxembourg Palace in Paris as they glittered in the light of the setting sun, and noticed that with the Iceland spar in a definite direction only *one* image was to be seen. He repeated his experiments with other sources of light, the light being reflected from glass plates or water. He concluded that light rays have definite " sides " and thought that this could be explained by the idea that the corpuscles radiated by the source of light were " polar ", i.e. had poles like little magnets. The name polarization is due to MALUS. FRESNEL established on the basis of the wave theory of light, which he strongly upheld, that two light rays polarized at right angles to one another cannot interfere with one another. The gifted physician THOMAS YOUNG (1773–1829) concluded from this that the vibrations of light must be transverse (see above) in their nature (1817). (Curiously enough, the idea that the vibrations of light are transverse had already been put forward by HOOKE in 1672.) In 1821 FRESNEL then found that transverse vibrations may be combined to form vibrations of very differing types, e.g. elliptic vibrations. He founded the theory of circular polarization (p. 252) and explained the rotation of the plane of polarization in quartz by means of it. He also explained double refraction as due to the wave surface not being spherical.

On p. 624 of Vol. III we learned that the electromagnetic radiation of a dipole is polarized, i.e. that the periodic variation of the electric intensity at some distance from the dipole takes place in one plane only; this was also proved by the experiments mentioned on p. 638 of the same volume. By analogy we should expect a priori that light should be polarized, if we are to explain light as consisting of the vibrations of an electrical configuration.

"Ordinary" light, however, is *not* polarized.

The reason for this lies in the structure of ordinary light, which we have already referred to repeatedly (pp. 2, 174). We must, however, infer from the fact that coherent light is capable of giving rise to interference even when the difference of path is great, that *the individual wave-trains of which ordinary light consists are strongly polarized.*

For, as was shown by FRESNEL, coherent wave-trains are not capable of interfering with one another unless they are polarized in

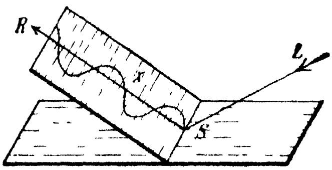


Fig. 7.—Light polarized in a plane at right angles to the plane of reflection; the vibrations take place in the plane of reflection

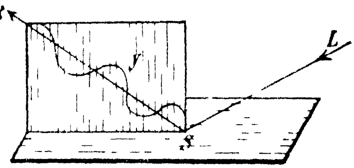


Fig. 8.—Light polarized in the plane of reflection; the vibrations take place at right angles to the plane of reflection

the same direction (§ 5, p. 241). If, however, this direction were to vary arbitrarily along the ray, lasting interference phenomena with large differences of path could not exist. The planes of polarization of the various individual wave-trains, on the other hand, are in quite arbitrary positions relative to one another, so long as the source of light is not under the special influence of an external electric or magnetic field. On the average, therefore, no plane has a privileged position, so that in general ordinary light is not polarized.

If a pencil of light rays L meets the plate of black glass P at the polarizing angle of 55°, only those rays are reflected which are vibrating in a single direction inclined at a perfectly definite angle to the plane of the incident light. It was long and vehemently disputed whether the plane in which the light is vibrating coincides with the plane of the incident and polarized rays, the plane of polarization (as shown in fig. 7), or is at right angles to it (as in fig. 8); but the question is seen to be pointless now that light is recognized as consisting of electromagnetic waves, in which the electric intensity vibrates (i.e. varies periodically) in one plane and the magnetic intensity in a plane at right angles to it. The two phenomena, the electric and

the magnetic, are of equal importance (Vol. III, pp. 616–617). Hence we can now understand how all optical phenomena may be represented by either assumption about the direction of vibration of light.

By carrying out experiments entirely analogous to those on the polarization of light (described above) with long electric waves, for which the position of the electric vector may at once be found (p. 231), it has been found that *the electric vector vibrates at right angles to the plane of polarization* as defined on p. 224, i.e. as in fig. 8. Hence *the magnetic vector vibrates in the plane of polarization*, as in fig. 7.

For historical reasons it has been agreed to call the direction in which the *electric* vector vibrates, i.e. the direction of the electric intensity E, the "direction of vibration of the light".

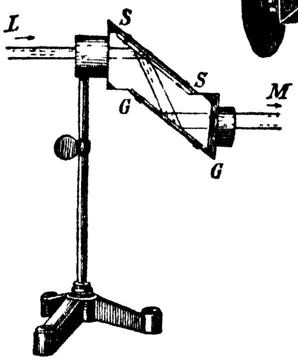


Fig. 9.—Grimsehl's reflection polarizer

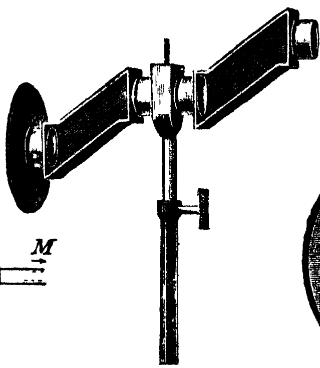


Fig. 10.—Two reflection polarizers capable of being rotated relative to one another

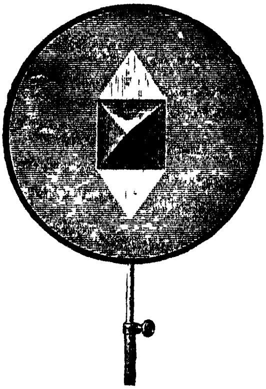


Fig. 11.—Reflection of polarized light at a four-sided glass pyramid

Thus if we say that light is polarized in the plane of incidence, or plane of the incident and reflected rays, this means that its electric vector vibrates at right angles to this plane.

Apparatus for Polarizing Light.—To avoid the inconvenience of apparatus in which the pencil of rays has its direction changed as a result of reflection, an ordinary silvered piece of glass is set parallel to the black reflecting plate and the whole enclosed in a solid case. We thus obtain the apparatus of fig. 9, which we may call a reflection polarizer. The pencil of parallel rays L coming from any source of light, e.g. the sun, is reflected by the mirror SS on to the glass plate GG. Here the reflected ray is again reflected in the direction M. Since SS and GG are parallel to one another, the reflected ray M is parallel to the incident ray L. The apparatus consisting of the mirror and glass plate, and hence the plane of incidence of the light, may be rotated at will in the ring of the support,

without the direction of the reflected ray M being altered. The mirror and plate are fixed into the case in such a way that the ray of light incident along the axis of the ring of the support makes an angle of 55° with the glass.

By combining two of these polarizers we obtain the apparatus shown in fig. 10. Here the second part of the combination of mirrors reflects the light polarized by the first part with maximum intensity when the planes of the incident and reflected rays in the two parts are parallel, whereas when these planes are at right angles the light is completely extinguished, the part of the light reflected in the first half of the apparatus being completely absorbed in the second half.

The behaviour of light polarized by reflection is brought out particularly clearly by the following apparatus.

A hollow square pyramid whose side faces are plates of black glass (fig. 11) is set with its base at the middle of a white screen in such a way that it can be rotated on its axis from behind. The angle of inclination of the side faces is chosen in such a way that the light incident in the direction of the axis meets the pyramid at an angle of 55°. If polarized light, produced by the polarizer in the position shown in fig. 9, is allowed to fall on this pyramid in the direction of the axis passing through the vertex, two bright triangular spots of light appear on the screen above and below the pyramid, which are produced by reflection of the light from the top and bottom surfaces of the pyramid, while the light is not reflected from its right and left surfaces. If the pyramid is rotated slowly about its axis, the bright spots naturally rotate also, and their intensity diminishes. At the same time the two side faces produce triangular spots, which are at first faint, but brighten as the rotation is continued. If the pyramid is in such a position that the diagonal of the base is vertical, i.e. that the planes of incidence of the side faces are inclined at an angle of 45° to the plane of incidence of the polarizer, all the four triangular spots due to reflection are of equal brightness. If the rotation is continued further, the spots which were formerly bright disappear, while the others continue to increase in brightness until their planes of reflection are parallel to those of the polarizer, that is, till the reflecting surfaces of the pyramid are at the top and bottom.

The glass pyramid shows in which planes the pencil of light falling on it is polarized. Any apparatus which serves this purpose is in general referred to as an **analyser.**

If the planes of polarization of the polarizer and the analyser are parallel to one another, we say that the polarizer and analyser are parallel, or we speak of "parallel Nicols" (see below, p. 239). If the planes of polarization of the polarizer and the analyser are at right angles to one another, we say that the polarizer and analyser are crossed, or we speak of "crossed Nicols".

An instrument formed by the combination of a polarizer and an analyser, as in fig. 10, is called a **polariscope.***

The intensity of the light which has passed through a polariscope is a maximum when the polarizer and the analyser are parallel, and a minimum when they are crossed. In the latter case the intensity is zero if the two parts of the polariscope polarize the light *completely*.

If the planes of polarization make an angle other than 0° or 90°

* A polariscope similar in principle to the apparatus of figs. 5 and 10, which was much used formerly, was constructed by J. G. C. Nörrenberg (1787–1862); he demonstrated his apparatus at a scientific congress at Karlsruhe in 1858.

with one another, the light polarized by the polarizer is only *partially* reflected at the analyser. According to MALUS, the following law holds:

The intensity of the light emerging from the analyser is proportional to the square of the cosine of the angle between the polarizer and the analyser.

If the intensity when the polarizer and the analyser are parallel is I and the angle between the polarizer and the analyser is ϕ, i, the intensity of the light leaving the analyser, is given by $i = I\cos^2\phi$.

The pyramidal analyser shown in fig. 11 splits up the polarized light into two components, whose amplitudes may be determined by the parallelogram law (Vol. II, p. 221; see also fig. 30 on p. 243 of the present volume). Hence the intensities of the two components

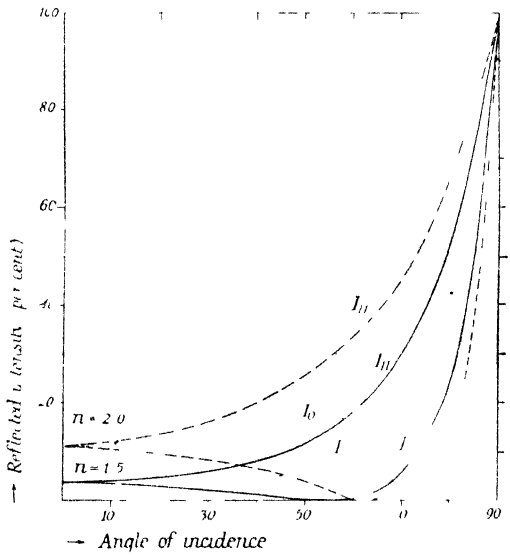


Fig. 12.—Variation of the intensity of the reflected light with the angle of incidence for light of various kinds

which are measured by the energy of the radiation and are proportional to the squares of the amplitudes (Vol. II, p. 221), are proportional to the squares of the cosine and of the sine of that angle through which the analyser has been rotated from its initial position, which corresponds to the maximum of reflection. Hence the sum of the intensities of the light reflected from two adjacent faces of the pyramid has always the same constant value.

The exact formulæ for the intensity and degree of polarization of the reflected ray in terms of the refractive index and the angle of incidence were given by FRESNEL. Fig. 12 illustrates the relation-

ships for transparent reflecting substances with refractive indices 1·5 and 2. The curves $I_{\parallel}$ are for light polarized in the plane of incidence,* the curves $I_{\perp}$ for light polarized at right angles to the plane of incidence, and the curves I_0 for ordinary unpolarized light.

2. Polarization by Refraction.

Polarization of light may arise from several other processes besides reflection.

If a parallel pencil of ordinary light is allowed to fall at the polarizing angle of 55° on a glass plate whose rear face is not blackened, part of the light is reflected in the manner described above; another portion of the light, however, passes through the glass plate after being twice refracted. This portion, as we may observe by means of an analyser, is also polarized. The portion of light passing through the glass plate, however, differs from the reflected portion in two respects:

(1) The polarization is not *complete*; on passing through the analyser the light is not *entirely* extinguished in any direction, only a trifling fluctuation of intensity being observed when the analyser is rotated.

(2) The plane of polarization of the transmitted light is at right angles to that of the reflected light; in the very position of the analyser in which the reflected ray is extinguished, the transmitted ray exhibits its maximum intensity, while in the position of the analyser in which the reflected ray exhibits its maximum intensity the transmitted ray exhibits its minimum intensity.

The polarization of light which occurs on refraction is explained by the fact that the light *reflected* at the boundary surface is polarized in the plane of incidence (that is, the electric vibrations are at right angles to the plane of incidence). Seeing that in ordinary light no plane of vibration occupies a special position, the remainder of the light, which is refracted into the transparent body, must be minus those vibrations which are present in the reflected light. The polarization associated with refraction is accordingly a sort of residual effect.

To obtain more complete polarization of the transmitted light, the light passing through one glass plate is made to pass through a second plate parallel to the first, through a third plate, and even more. A layer of glass consisting of about twenty flat pieces of plate glass with only thin spaces between them will polarize the light transmitted by it almost completely, provided the pencil of light rays is incident at the polarizing angle of 55°. Such a combination of a large number of pieces of plate glass to form a single instrument is called a **pile of plates**. This may be used either as a polarizer or as an analyser.

Fig. 13 shows how a pencil of rays L behaves on falling on a pile of plates

* That is to say, the electric vector is vibrating at right angles to the plane of incidence.

at the polarizing angle of 55°. It is split up into the reflected part M and the refracted part N, which leaves the pile of plates in a direction parallel to L. In the ray M, which is polarized by reflection, the electric vector vibrates at right angles to the plane of incidence. (The plane of incidence coincides with the plane of the paper; the points marked on the ray M are intended to indicate that the electric vector is at right angles to the plane of the paper.) The transmitted ray N polarized by a pile of plates, and the reflected ray, are polarized in directions at right angles; the direction of vibration of the electric vector lies in the plane of incidence. (The short cross dashes on the ray N are intended to indicate that the electric vector is in the plane of the figure.) The ray M polarized by reflection is reflected to a small extent towards R by the second pile of plates A_1, which acts as an analyser; the rest of the energy of this ray (85 per cent) enters the glass. The ray N polarized by refraction passes through the pile of plates A_2, which acts as an analyser, in the direction B without any reflection, provided the pile of plates P is so thick that the ray N may be regarded as completely polarized.

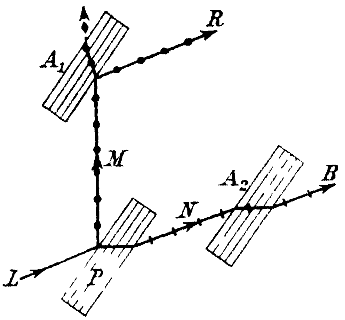


Fig 13.—The reflected ray and the transmitted ray are polarized in directions at right angles to one another

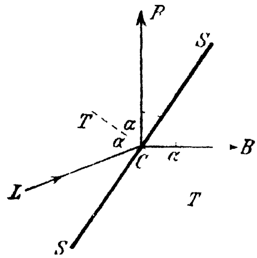


Fig. 14.—In reflection at the polarizing angle the reflected ray and the refracted ray are at right angles to one another

Brewster's Law.—The peculiar fact that polarization by reflection (and also by refraction) takes place most completely at a perfectly definite angle (55° for glass) varying for different substances makes it natural to suppose that there is a simple relationship between the refractive index of a substance and its polarizing angle. The relationship was discovered experimentally in 1813 by Brewster (see p. 187), who also found a theoretical basis for it. Brewster's law is as follows:

The polarizing angle is determined by the fact that the refracted and reflected parts of the incident ray are at right angles to one another.

Then if the ray of light L (fig. 14) falls at the polarizing angle α on the boundary SS of the refracting substance of refractive index μ'/μ, the reflected ray CR must be at right angles to the refracted ray. If this is the case the polarizing angle α and the angle of refraction α' corresponding to it are connected by two equations, namely, (1) the law of refraction $\mu \sin\alpha = \mu' \sin\alpha'$ and (2) the relationship

$\alpha + \alpha' = 90°$, which is obvious from the figure. From these two equations it follows that

$$\sin\alpha = \frac{\mu'}{\mu}\sin(90° - \alpha) = \frac{\mu'}{\mu}\cos\alpha,$$

whence $$\tan\alpha = \frac{\mu'}{\mu}.$$

Determination of the Constituents of the Light-vector.—In order to settle the question in which plane the electric vector vibrates and in which the magnetic vector, KLEMENČIČ * placed a pile of plates consisting of a number of thick glass plates in the path of electric waves in such a way that the electric waves fell on the pile of plates at the polarizing angle. When he investigated the intensity of the transmitted rays as he rotated the pile of plates about the incident ray, he found that the intensity of the rays is greatly weakened when the electric vector is at right angles to the plane of incidence of the electric waves, while the electric waves pass through the pile of plates with almost undiminished intensity when the electric vector is parallel to the plane of incidence. From this follows the result which we assumed in advance on p. 226:

In light rays polarized by reflection the electric vector vibrates at right angles to the plane of incidence; in the transmitted polarized light the electric vector vibrates parallel to the plane of incidence to the extent corresponding to the reflected intensity.

For the detection of the polarization of light rays by metal gratings see Vol. III, pp. 643–644.

Note.—The above remarks on polarization by reflection and refraction hold good for transparent substances only, i.e. for regions of wave-length where there is no marked absorption.

3. Polarization associated with Diffraction. The Tyndall Effect.

Even ARAGO noticed the fact that light is partially polarized on *diffraction*. For example, polarized light diffracted by a grating in general has its plane of polarization altered. As DU BOIS and RUBENS established by means of wire gratings (1904), rays whose electric vector is vibrating parallel to the wires are more markedly weakened than rays whose electric vector is vibrating at right angles to the plane of the grating (Vol. III, p. 644). The analogous experiments of HERTZ were likewise discussed on p. 638 of Vol. III. For these long waves the transparency or opacity of the grating is actually perfect.

* J. KLEMENČIČ (1853–1901), Professor of Physics at Graz and Innsbruck.

The polarization phenomena observed in connexion with the *scattering* of light are a mixture of polarization by reflection and polarization by diffraction.

If light passes through a turbid medium (diluted milk, an alcoholic solution of mastic diluted with water, a smoke of fine particles, &c.), the light is diffracted sideways and scattered, so that the path of the rays may become visible (the so-called *Tyndall * effect*). LORD RAYLEIGH † has shown (1871, 1899) that I, *the intensity of the scattered light, is inversely proportional to the fourth power of the wave-length* λ, *provided the scattering particles are small compared to* λ:

$$I = \frac{C}{\lambda^4}.$$

If the originally incident light is white, the scattered light will accordingly contain much more blue than red; that is, the scattered light will appear bluish and the transmitted light reddish.

It has been found that all substances, no matter how carefully purified, are "turbid"; this is to be ascribed to the molecular structure of matter. Under certain circumstances the wave-length of the incident light is slightly altered in the scattering process (the *Raman effect*). For further details see Vol. V.

The light diffracted by a turbid medium is **polarized.** Thus if a small quantity e.g. of an alcoholic solution of mastic is shaken into a long glass trough filled with water and a parallel pencil of bright white light is sent through the trough, the turbid mixture shines with a blue light in a direction at right angles to the path of the rays.

If the scattered light is observed by means of an analyser, we find that if we look in a direction at right angles to the ray of light the scattered light is almost completely polarized in the plane passing through the ray and the line of observation. In turbid media containing insulating particles the maximum polarization occurs in this direction, i.e. at an angle of 90° to the ray, but in turbid media containing metallic particles (Ag, Au, Pt) it occurs at an angle of 110°–120° to the ray (EHRENHAFT).

* JOHN TYNDALL (1820–93), studied from 1848–50 at Marburg (under Bunsen) and in 1851 at Berlin (under Magnus); from 1853 onwards he was Professor of Physics at the Royal Institution and at the School of Mines in London as the successor of MICHAEL FARADAY. He was celebrated in his day on account of his brilliant skill as an experimenter and his fascinatingly written popular works upon various physical subjects; he was also a prominent supporter of mountaineering sport in the Alps.

† JOHN WILLIAM STRUTT, third BARON RAYLEIGH (1842–1919), Cavendish Professor of Experimental Physics in the University of Cambridge from 1879 to 1884, and Professor of Natural Philosophy at the Royal Institution of Great Britain from 1887 to 1905.

4. Double Refraction.*

Iceland spar (calcite, calcspar) is a colourless transparent substance crystallizing in the hexagonal (rhombohedral) system; chemically it consists of $CaCO_3$, and particularly fine specimens are found in Iceland (hence the name Iceland spar). The most common form of crystal is shown in fig. 15. If a piece of Iceland spar of any shape is struck with a hammer, it exhibits a remarkable tendency to cleave obliquely in three definite directions, the fragments formed by cleavage being rhombohedra.

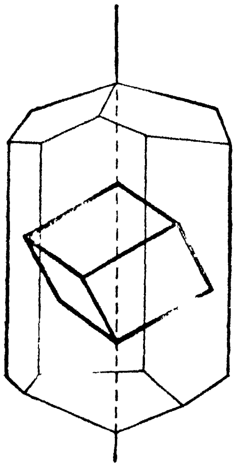


Fig. 15.—An ordinary crystal of Iceland spar with inscribed rhombohedron

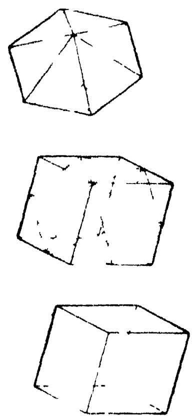


Fig. 16.—The rhombohedron is the hemihedral form of the six-sided pyramid

In fig. 15 a rhombohedron produced by cleavage is drawn in the correct position relative to the natural crystal. The rhombohedron is bounded by six faces, each of which is a rhombus with an angle of 102°, and each of which is inclined to the next at an angle of 105° 6′. At two opposite corners of the rhombohedron all the angles of the faces meeting there are obtuse, while at the other six corners there are always one obtuse angle and two acute angles. Fig. 16 shows how a rhombohedron of Iceland spar may be regarded as a hemihedral form of a six-sided double pyramid.

The line joining the two blunt corners coincides with the principal crystallographic axis of the Iceland spar, which is also the optic axis, or line along which the crystal is singly refracting (p. 238). Every plane passing through the principal axis is called a *principal section*.

If we lay a rhombohedron of Iceland spar on a printed page we see the type double (fig. 17).

* Double refraction was discovered in 1669 by the Dane ERASMUS BARTHOLINUS (1625–98), who was a mathematician, a doctor, and a lawyer; it was investigated in detail by HUYGENS.

If a rhombohedron of Iceland spar is brought into the path of a narrow pencil of parallel rays (L in fig. 18) in such a way that the latter falls perpendicularly on the front surface of the rhombohedron at I, the light leaving the Iceland spar is split up into two parallel pencils. One of these pencils (O) goes through the Iceland spar without refraction, while the other (E), in spite of the normal incidence of the original pencil, is deviated on entering the Iceland spar, and is deviated to the same extent in the opposite direction on leaving it, so that the two separate pencils leave the Iceland spar parallel. The first pencil (O), which behaves like any pencil falling normally on a plate of glass bounded by parallel planes, i.e. obeys the ordinary laws of refraction, is called the **ordinary ray**; the second pencil (E), which

Fig. 17.—Double refraction of Iceland spar

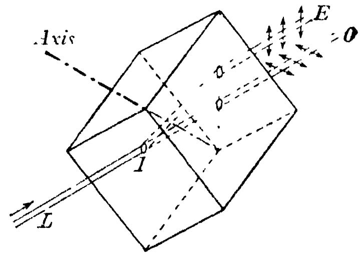


Fig. 18.—Ordinary ray and extraordinary ray

owing to its double deviation is displaced sideways relative to its original position, is called the **extraordinary ray.** If a rhombohedron of Iceland spar is rotated about the normally incident ray LI, the ordinary ray remains in the same position, while the extraordinary ray rotates in a circle about the ordinary ray.

The refraction and displacement of the extraordinary ray in Iceland spar always take place in the plane of the principal section, and in such a way that the extraordinary ray makes a greater angle with the optic axis than the ordinary ray does. Hence, on emerging from the rhombohedron, the extraordinary ray is farther from the blunt corner of the rhombohedron than the ordinary ray is. The angle between the ordinary ray and the extraordinary ray in Iceland spar is always the same when the light is incident normally. Hence the thicker the rhombohedron, the greater the distance between the ordinary ray and the extraordinary ray. The intensities of the two pencils leaving the Iceland spar are exactly the same.

If we investigate the two pencils with an analyser, we find that the two rays are *completely* polarized in directions at right angles to one another.

The plane of polarization of the ordinary ray coincides with the principal section, that of the extraordinary ray being at right angles to the principal section.

The arrows drawn across O and E (fig. 18) are meant to indicate the directions in which the electric vector is vibrating in the two rays.

The ordinary ray is said to be polarized in the principal section, the extraordinary ray at right angles to the principal section.

If the two pencils are made to pass through an analyser, and the latter is rotated, the brightness of the two rays varies in the same way as the brightness of the triangular spots did in the experiment with the pyramidal analyser (fig. 11, p. 226).

If the two pencils are made to pass through a second rhombohedron of Iceland spar the same as the first, no further splitting-up takes place, provided the principal axes of the two rhombohedra are

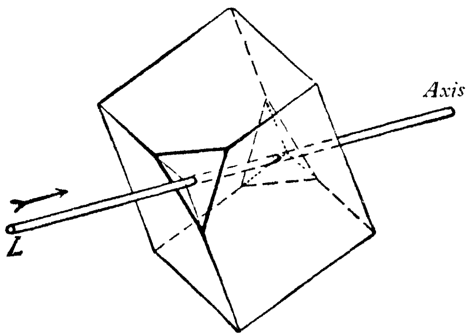


Fig. 19.—No double refraction takes place in the direction of the short diagonal of the rhombohedron.

parallel. If the second rhombohedron is rotated through 180° relative to the first about the ordinary ray, the displacement of the extraordinary ray in the second rhombohedron is in the opposite direction to that in the first, and the two rays emerge recombined from the second piece of Iceland spar as ordinary unpolarized light. If the second rhombohedron is rotated through 90° about the incident ray, the extraordinary ray of the first rhombohedron passes through the second rhombohedron as the ordinary ray and the ordinary ray of the first rhombohedron as the extraordinary ray.

If a parallel pencil of rays falls *obliquely* on the front surface of the rhombohedron of Iceland spar, the pencil is likewise split up into two parts, but *both* parts are now refracted. The refractive index for the two rays may be determined from the angles of incidence and refraction. The following results are obtained.

For every angle of incidence the refractive index for the ordinary ray is 1·65 (*for the D line* 589 mμ). *The refractive index for the extraordinary ray, on the other hand, alters with the angle of incidence, varying between* 1·48 *and* 1·65.

The refractive index for the extraordinary ray has its maximum value (1·65) when this ray traverses the Iceland spar parallel to its

principal axis, and its minimum (1·48) when it traverses the Iceland spar in a direction at right angles to the optic axis.

If we cut off the blunt corners of a rhombohedron of Iceland spar in such a way that the boundary planes are at right angles to the optic axis (fig. 19), a ray falling normally on this plate, i.e. along the optic axis, is not split up.

Wave Surfaces in Doubly-refracting Crystals.—As the refractive index for the ordinary ray has a constant value (1·65) independent of the angle of incidence, we may assume that the velocity of light for the ordinary ray is the same in all directions, since the refractive index is given by the ratio of the wave velocities of light in air and in the refracting substance.

As, on the other hand, the refractive index for the extraordinary ray varies with the angle of incidence, the wave velocity of light for the extraordinary ray must vary in different directions. It is a minimum in the direction of the optic axis and a maximum in a plane at right angles to this axis. Within this plane, however, the velocity of propagation of light is the same in all directions.

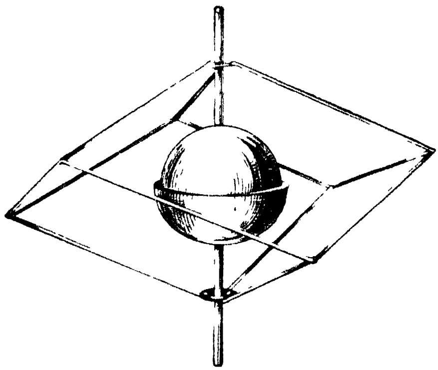


Fig. 20.—Wave surfaces of the ordinary ray (upper hemisphere) and the extraordinary ray (lower hemi-ellipsoid) in Iceland spar

Following Fresnel's method (1821), we may throw light on this behaviour by means of the following conception. In the interior of a large piece of Iceland spar we imagine a single point acting as the centre or source of an optical wave system. The light waves are propagated in two ways from this point as centre. The part associated with the ordinary ray, for which the velocity of propagation is the same in all directions, after a very brief interval of time, say 10^{-6} second, has traversed the same distance in all directions; it has therefore reached the surface of a sphere with its centre at the source of the light waves. The second part, corresponding to the extraordinary ray, has different velocities in different directions. In the direction of the optic axis the velocity of the extraordinary ray is the same as that of the ordinary ray. In the direction at right angles to this, however, the velocity of the extraordinary ray is greater in the ratio of 1·65 to 1·48, as in this direction the refractive index is only 1·48. Corresponding calculations for other directions, based on the values of the refractive index for the extraordinary ray, show that the extraordinary ray reaches the surface of an ellipsoid of revolution which is shortened along the optic axis. Its axis of revolution coincides with the principal axis of the Iceland spar. The resulting combination of a sphere and an ellipsoid of revolution is called Fresnel's wave surface; in fig. 20 it is shown in its correct position relative to a rhombohedron of Iceland spar. The upper part (a hemisphere) represents the wave surface for the ordinary ray, the lower part (half of an ellipsoid of revolution) represents the wave surface for the extraordinary ray. We have to imagine both halves completed, so that the ellipsoid of revolution completely encloses the sphere.

Application of Huygens' Principle.—Let ZZ (fig. 21) be the plane boundary surface of a piece of Iceland spar. Let the space above ZZ be occupied by air,

the space below ZZ by Iceland spar. A pencil of parallel rays travelling through the air and incident on the plate of Iceland spar in the direction I meets the boundary surface at A at a time when the remoter part of the pencil of rays

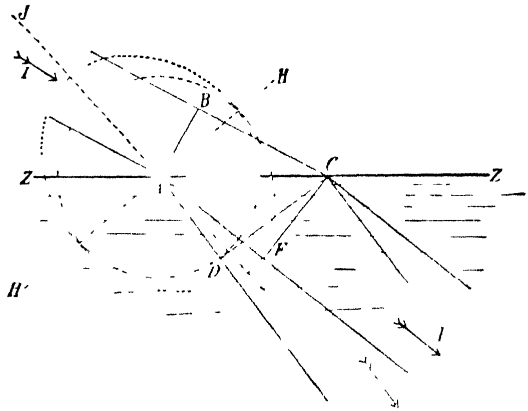


Fig. 21.—The directions of the ordinary ray and the extraordinary ray according to Huygens' principle

is still at B, i.e. has not yet reached the surface. A becomes the centre of two elementary waves, which traverse the Iceland spar with differing velocities, one of them, corresponding to the ordinary ray, spreading out spherically, and the other as an ellipsoid. Let the dotted line HH represent the optic axis of

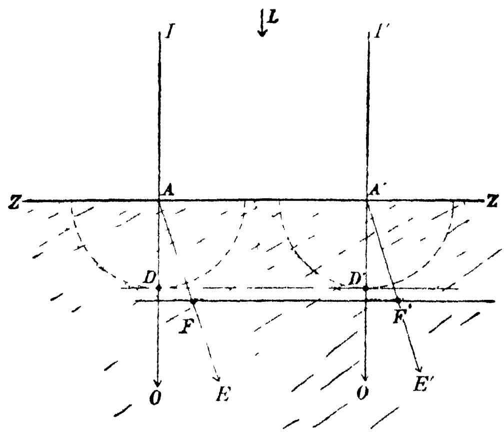


Fig. 22.—Double refraction of light incident on the surface of the crystal at right angles

the Iceland spar. When the part of the pencil of rays corresponding to the point B reaches the boundary surface at C, the wave surface associated with the ordinary ray from A has spread out into a sphere, which is indicated in the figure by the dotted circle. The ratio of its radius to the distance BC, i.e. the ratio of the velocity of light along the ordinary ray in Iceland spar and the velocity of light along the incident ray in air, is 1 : 1·65. The wave-front corresponding to the

sphere is found by drawing CD, the tangent plane from C to the sphere. The direction of the ordinary ray, which is indicated by the arrow O, is at right angles to CD.

The elementary wave corresponding to the extraordinary ray from A is represented by the dotted ellipse, whose major axis JJ is at right angles to HH, the optic axis of the Iceland spar. The wave-front corresponding to the extraordinary ray is found by drawing CF, the tangent plane from C to the ellipsoid of revolution. The direction of the extraordinary ray, which is indicated by the arrow E, is *not* at right angles to the wave-front CF. In the case of the extraordinary ray, therefore, it is necessary to distinguish between a *ray velocity* in the direction of E and a *normal velocity* at right angles to the wave surface CF.

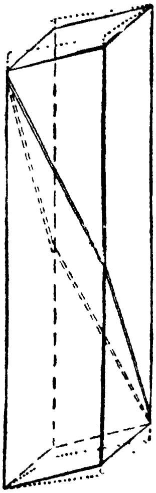


Fig. 23.—Nicol's prism

It follows from the figure that if the angle of incidence of the ray E is altered the refractive index for the ordinary ray remains unaltered, as the ratio AD : BC remains the same. For the extraordinary ray, on the other hand, this ratio alters with the angle of incidence, as CF, the tangent to the ellipse, takes up a different position for every value of the angle of incidence. Hence it follows that the ratio AF : BC also varies with the angle of incidence. This ratio reaches its maximum value 1 : 1·48 when the point of contact of the tangent lies in the direction AJ, i.e. when the ray travels through the Iceland spar at right angles to the optic axis. In this case the angle between the directions of the rays O and E is a maximum.

If the extraordinary ray travels through the Iceland spar parallel to the optic axis, AF assumes its least possible value, i.e. the refractive index for the extraordinary ray then reaches its greatest possible value, namely, that for the ordinary ray. A ray of light which traverses the Iceland spar in the direction of the optic axis is not split up.

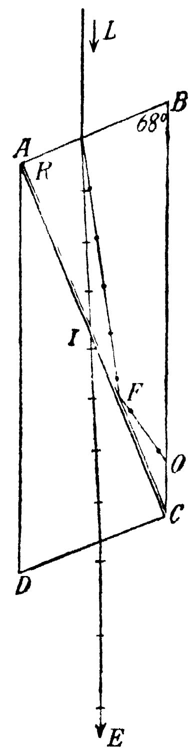


Fig. 24.—Paths of the rays in the Nicol prism

Some interest attaches to the particular case where the incident pencil of light falls normally on the surface of the Iceland spar, as in fig. 22. Let IA, I′A′ represent the boundaries of the pencil incident normally along L on the boundary surface. From A and A′ the elementary waves spread out simultaneously in all directions. The two wave surfaces are drawn round A and A′. DD′, the common tangent plane of the spheres, is the wave-front for the ordinary ray, which advances parallel to AO. The wave-front for the extraordinary ray is determined by FF′, the common tangent plane of all the ellipsoids. This fixes the direction of the extraordinary ray, which is parallel to AE.

This figure brings out very clearly why the direction of the extraordinary ray AE is no longer at right angles to the wave-front. This is why the extraordinary ray does not obey the law of refraction, which was deduced from Huygens' principle on p. 249 of Vol. II on the assumption that the direction of the ray is at right angles to the wave-front.

Nicol's Prism.—The two rays which emerge from a piece of Iceland spar in which one ray of ordinary light has been split up are *completely* polarized in

directions at right angles to one another. If it were possible so to separate the ordinary ray and the extraordinary ray in space that the rays could be worked with separately, Iceland spar would be particularly suitable for use as a polarizer.

The rays can be completely separated by an artifice used by W. NICOL * in 1841. The ends of a longish rhombohedron of Iceland spar produced by cleavage are ground away (as shown by dotted lines in fig. 23) until they make an angle of only 68° (instead of 71° as in the crystal obtained by cleavage) with the longitudinal edges. The piece of Iceland spar so treated is then cut through diagonally by a plane at right angles to the new end surfaces and at right angles to the principal axis, which crosses the Iceland spar diagonally. After the diagonal surfaces are ground flat and polished, they are cemented together by a thin layer of Canada balsam so as to occupy exactly the positions which they occupied previously.

Fig. 24 shows a diagonal principal section through the Nicol prism. If a pencil of rays L falls on the artificial end surface AB in a direction parallel to the unaltered longitudinal edges, the pencil is refracted and split up into an ordinary ray and an extraordinary ray, the ordinary ray being more strongly refracted than the extraordinary ray. The refractive index of Canada balsam is less than that of Iceland spar. Hence total reflection can occur at the layer of Canada balsam, provided the light is incident at an angle exceeding the critical angle for total reflection. Now the angles are so calculated that the ordinary ray falls on the Canada balsam at F at an angle exceeding the critical angle for total reflection and as a result is reflected to one side out of the prism. in the direction of O. The angle of incidence of the extraordinary ray is less than the critical angle for total reflection; hence it passes through the layer of Canada balsam at I and finally leaves the Nicol prism in the same direction (E) as it met it, with a trifling lateral displacement. The Nicol prism is the most perfect form of polarizer, as apart from this trifling lateral displacement and the desired polarization the light is not altered in any disadvantageous way (e.g. its colour is unaffected). The intensity of the polarized light is of course less than that of the incident light, as the other polarized portion of the light has been removed and the light is also weakened by reflection at the boundary surfaces.

Polarization by Tourmaline Plates.—When an ordinary non-polarized ray of light traverses a homogeneous isotropic medium, there is nothing to cause the light to have any tendency to vibrate in any one direction rather than in another. Hence in this case there is no polarization of the light, nor is there any when the light falls vertically on the boundary surface of two transparent media. The circumstances are different, however, when a ray of light enters a crystal whose internal structure, like e.g. that of Iceland spar, differs in different directions. In fact, many minerals have the property of polarizing the light which they transmit. This phenomenon is exhibited in a striking way by tourmaline, which crystallizes in the hexagonal (rhombohedral) system, and whose permanent electrical polarization we have already mentioned in Vol. III (p. 95).

If a thin plate bounded by parallel planes is cut from a brown or green crystal of tourmaline, the parallel planes being parallel to the principal crystallographic axis (i.e. parallel to the plane of the paper in fig. 118 (Vol. III, p. 95)) and the plate being of such a thick-

* W. NICOL (1768–1851), a lecturer on physics in Edinburgh.

ness that the transmitted light is coloured, it is found that light incident normally on the tourmaline plate is polarized. If we lay two tourmaline crystals one on top of the other so that the axes XX of the

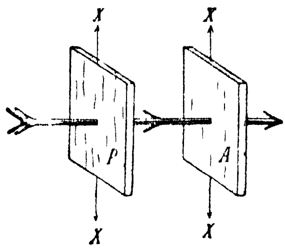


Fig. 25.—Parallel plates of tourmaline

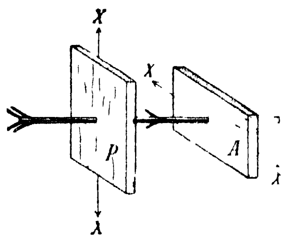


Fig. 26.—Crossed plates of tourmaline

crystals are parallel (fig. 25), the light passes through both plates. If, however, we turn one of the tourmaline plates relative to the other, the intensity of the light steadily diminishes; when the axes are at right angles (crossed tourmalines, fig. 26) no light is transmitted.

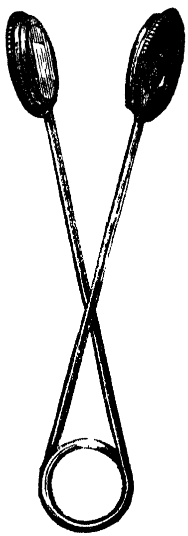


Fig. 27.—Tourmaline forceps

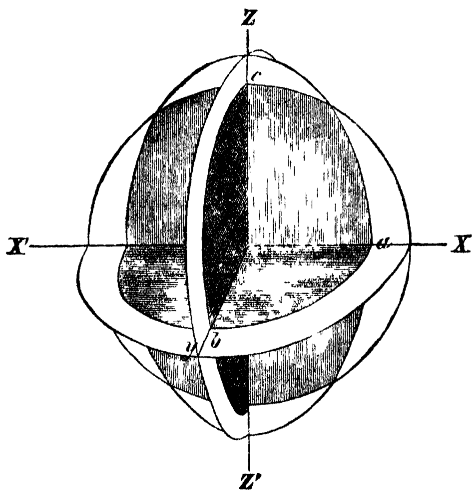


Fig. 28.—Wave surface for biaxial crystals

Two tourmaline plates are usually combined to form a polarizer in the way shown in fig. 27. The tourmaline plates are mounted in circular rings which may be rotated relative to one another and may be pressed together by means of a spring of bent wire, so that a substance under investigation can be held tightly between the two tourmaline plates. This type of polarizer is called a **tourmaline forceps**.

Note.—In a polarizing apparatus any two of the following may be combined in any way: reflecting polarizer, pile of plates, Nicol prism, and tourmaline plate. There is no essential difference between polarizer and analyser. In general the part of the apparatus on which ordinary non-polarized light falls and in which it is polarized is called the polarizer, and the part by which the polarized light is observed is called the analyser.

The polarizing effect of the tourmaline plate is due to the fact that, like Iceland spar, it splits the light into an ordinary ray and an extraordinary ray. Even in quite small thicknesses, however, tourmaline absorbs the ordinary ray completely, so that the extraordinary ray—of course much enfeebled—alone leaves the plate. The property of absorbing the ordinary ray and the extraordinary ray to different extents is called **dichroism,*** as owing to the differing absorption of the crystal for different directions of vibration of the transmitted light (and also, as a rule, for different wave-lengths) the crystal if sufficiently thin appears to vary in colour when looked at in different directions.

Uniaxial and Biaxial Crystals.—It is only crystals of the regular system and amorphous substances that are singly refracting. All crystals in which *two* of the crystallographic axes are different, i.e. which crystallize in the tetragonal or hexagonal system, behave like Iceland spar. If the ellipsoid of revolution of the Fresnel wave surface is elongated in the direction of the optic axis, the crystal is said to be **positive,** while a crystal which behaves like Iceland spar is said to be **negative.**

Non-cubic crystals which do not crystallize either in the tetragonal or in the hexagonal system, i.e. which have *three* different crystallographic axes, likewise exhibit double refraction, but the wave surface representing the double refraction is of the form shown in fig. 28. In these crystals the velocity of propagation of the light is the same for the two polarized components of the light in two definite directions. These two directions of equal velocity of propagation are called **optic axes** and the crystal is said to be **biaxial.** By optic axes we mean the directions of equal *normal* velocities, which we obtain for the *normal surface* by the same construction as that carried out on p. 237 for the wave surface.

5. Interference of Polarized Light.

If, as with Iceland spar, a pencil of light rays is split up into two pencils polarized in directions at right angles to one another, the crystal plate produces a difference of phase between them, owing to the difference in the rate of propagation of light in the two pencils. By making the plate of Iceland spar of a suitable thickness, it is possible to arrange that the ordinary ray is half a wave-length in advance of the extraordinary ray on leaving the crystal. Nevertheless, interference bands or interference rings resembling diffraction phenomena or Newton's rings are never exhibited.

It follows that ***two rays of light polarized in directions at right angles to one another do not interfere, and can cross one another without disturbance.***

FRESNEL and ARAGO had established that two rays polarized at right angles to one another will not interfere under any circumstances, THOMAS YOUNG (1817) was the first to draw the conclusion that the vibrations of light must be transverse. Two rays of light polarized

* Gr., *dichroos*, two-coloured.

in directions at right angles to one another vibrate in two planes at right angles to one another; the vibrations are then entirely independent of each other and cannot interfere. If, however, the two components are brought into the same plane after passing through the Iceland spar, interference phenomena will arise. Such phenomena occur, e.g. if a thin plate of gypsum is inserted between the polarizer and the analyser.

A parallel pencil of monochromatic light (p. 4) is allowed to pass through a polarizer and an analyser in such a way that the light is extinguished; that is, the polarizer and the analyser are crossed (p. 227). If we bring a thin plate of gypsum into the space between the polarizer and the analyser, light will as a rule pass through. If we rotate the plate of gypsum, we obtain maximum intensity in a definite position. If we then rotate the analyser, the field of view is darkened again.

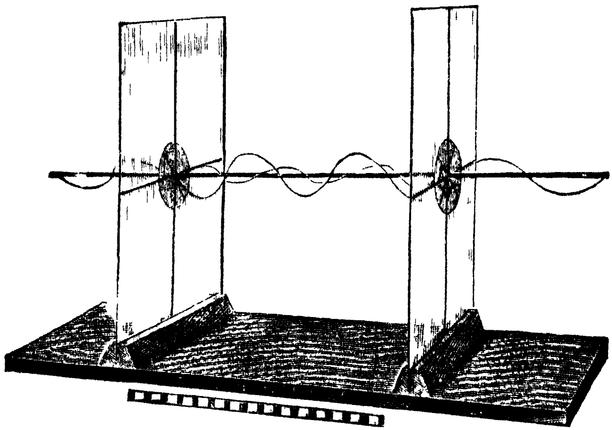


Fig. 29.—Model to illustrate the action of a thin plate of gypsum

The phenomena may be illustrated by the model shown in fig. 29. Two rectangular glass plates are set up vertically on a board; they represent the front and back surfaces of the plate of gypsum, so that the distance between them represents the thickness of the gypsum. On each plate a set of axes is drawn passing through the centre of the plate. Let the vertical axis represent the plane of vibration of the first ray and the horizontal axis that of the second ray into which the plate of gypsum would split up a ray of light entering it at right angles to the boundary surface. A bar whose three parts can be rotated relative to one another passes through holes in the two glass plates. The incident light is polarized by a polarizer on the left of both glass plates (not shown in the figure) in such a way that the plane of vibration meets the axis on the left-hand plate at an angle of 45° at the origin.

The polarized ray is split up into two components (Vol. II, p. 221) whose amplitudes may be constructed by the parallelogram law as in fig. 30. The resolution is shown in the model on the circular discs at the centre of the two glass plates. The components have the same amplitude provided that, as we

have assumed, the plane of vibration of the incident light makes an angle of 45° with the planes of vibration in the plate of gypsum. For other angles the amplitude of the component is given by $a \sin \varphi$ for the first ray O and by $a \cos \varphi$ for the second ray E, where a denotes the amplitude of the incident ray and φ the angle at which P, the plane of vibration of the incident light, is inclined to E, the plane of vibration of the second ray.

The two velocities of propagation, i.e. the wave-lengths of the two components, are different in the gypsum. In the model (fig. 29) it is assumed that the first ray covers $2\frac{1}{2}$ wave-lengths within the gypsum and the second ray 3 wave-lengths. The phase difference of the two rays is then half a wave-length (**half-wave plate**). In order to take account of the difference of phase, the amplitude $a \sin \varphi$ has been given the opposite sign in figs. 31 and 32, which refer to the rays emerging from the plate of gypsum, to what it has in fig. 30.

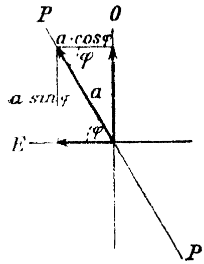


Fig. 30.—Resolution of the amplitude a into two components

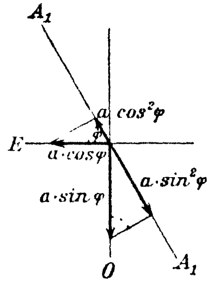


Fig. 31.—The components resolved again by the analyser

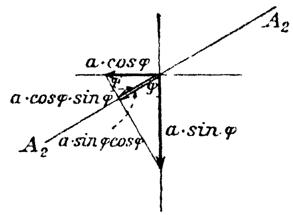


Fig. 32.—The components resolved again, the analyser being turned through 90°

When the ray, consisting of the two components, leaves the gypsum, nothing particular is to be observed, as the two components are polarized at right angles to one another. If the ray then passes through an analyser A_1A_1 (fig. 31) whose plane of vibration coincides with that of the polarizer, each component is separately split up into two new components by the analyser, and of these only those vibrating in the original plane, with the amplitudes $a \cos^2 \varphi$ and $a \sin^2 \varphi$, can pass through, the two other components being absorbed in the analyser. The two transmitted components, however, have a phase difference of half a wave-length (by fig. 31 they are in opposite directions) and if the amplitudes are equal ($\varphi = 45°$) they completely cancel one another.

If we then turn the analyser through 90° the two other components, with the amplitudes $a \cos \varphi \sin \varphi$ and $a \sin \varphi \cos \varphi$ (fig. 32), whose planes of vibration are at right angles to those just described, can alone pass through the analyser; and as these do not differ in phase (in fig. 32 they are in the same direction), their effects are additive, i.e. when $\varphi = 45°$ the light leaves the analyser with the amplitude $2a \sin \varphi \cos \varphi = a \sin 2\varphi = a \sin 90° = a$, i.e. with the intensity it would have if it passed through parallel polarizer and analyser and no gypsum. The difference in phase due to the gypsum has the effect that for $\varphi = 45°$ the light passes through without diminution when the polarizer and analyser are crossed, i.e. in the position where there would be complete darkness if there were no gypsum present.

This effect of gypsum is not fully exhibited unless the thickness of the gypsum is so adjusted that the difference of phase is exactly half a wave-length or an odd multiple of half a wave-length. If, on the

other hand, the thickness of the gypsum corresponds to exactly a wave-length or an integral number of wave-lengths, i.e. an even multiple of half a wave-length, the light goes through when the polarizer and analyser are parallel just as if the gypsum were not there, and when the polarizer and analyser are crossed the light is extinguished just as if the gypsum were not there.

Colour Phenomena with Polarized Light.—If instead of monochromatic light we allow white light to fall on the polariscope and the gypsum, the phase difference is an odd multiple of half a wave-length for certain definite colours and an even multiple for certain other colours, the colours depending on the thickness of the gypsum. Thus when the polarizer and the analyser are crossed, a perfectly definite portion of the light, e.g. the green rays, goes through unhindered; on the other hand, rays of slightly different wave-length are more or less enfeebled and part of the light is more or less extinguished, some of it completely. The gypsum then appears to have the colour of the transmitted light, in our case green.

If the analyser is then turned through 90°, those very parts which were transmitted before are now extinguished, while the parts which were previously extinguished are now transmitted; hence the gypsum now appears to have the complementary colour, in our case purplish-red. If the gypsum is rotated between the polarizer and analyser, the tint of its *colour* does not vary, but the *intensity* of its colour does, the depth of the colour becoming less, as the direction of vibration in the gypsum no longer makes an angle of exactly 45° with the direction of vibration in the polarizer and the analyser.

If we rotate the analyser or the polarizer without moving the gypsum, the one colour gradually gives place to its complementary colour.

If we vary the thickness of the gypsum the nature of the colour is changed. A plate of gypsum consisting of portions of differing thicknesses exhibits a great variety of colours between polarizer and analyser; hence by suitable arrangement of gypsum plates of different thicknesses it is possible to make beautifully coloured pictures (e.g. of butterflies).

Any thin sheet of a doubly-refracting substance when placed between polarizer and analyser will behave in the same way as gypsum. The colour phenomena may of course be exhibited to a number of people at once by letting the light transmitted by the apparatus fall on a white screen.

6. Convergent Polarized Light. Double Refraction arising from Strain.

If the plate of gypsum between polarizer and analyser is rotated about an axis at right angles to the ray of light, the path of the ray in the gypsum becomes longer the greater the angle through which

the plate of gypsum has been rotated out of its normal position. Hence it follows that the rotation alters the phase difference within the gypsum, i.e. the colour of the light transmitted by it.

The various rays of a convergent pencil of incident light, which subsequently diverges from the front boundary surface of a plate of some doubly-refracting substance, traverse the plate at various angles, i.e. have paths of varying length within the substance; hence the plate will not give rise to the same colour all over, but to rings varying more or less in colour. The simplest case arises when we bring a section

Fig. 33.—Plate of a uniaxial crystal between parallel Nicols

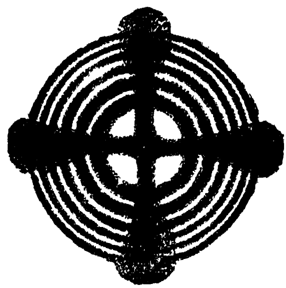


Fig. 34.—Plate of a uniaxial crystal between crossed Nicols

of a uniaxial crystal, e.g. Iceland spar, cut at right angles to the optic axis, into the path of a pencil of polarized light and investigate the transmitted light by means of an analyser. Here we obtain the ring system shown in fig. 33 when the polarizer and analyser are parallel and that shown in fig. 34 when the polarizer and analyser are crossed, a white or black cross being superposed on the rings. One limb of the cross coincides with the plane of polarization of the incident light, the other is at right angles to it. If the light used is monochromatic, the ring system consists of bright and dark rings; if it is not monochromatic, e.g. if white light is used, the rings are coloured, like Newton's rings.

We base our explanation of the phenomena on the model shown in fig. 35. The drawing must be imagined as three-dimensional. Let the rectangle surrounding the figure represent the boundary of a vertical plane, e.g. a frame of four bars. The vertical bars are crossed by the horizontal axis DD, which at the same time represents the optic axis of a plate of Iceland spar cut at right angles to the axis. The oblique parallelogram in the centre is to denote the front boundary surface of the plate of Iceland spar; the Iceland spar is supposed to be on the left, air on the right. In the plane indicated by the semicircle MAM′ let the polarized ray IC fall on the centre of the plate. Let the plane in which the light is vibrating be vertical; it is indicated by the small vertical arrow at I. Let the radius of the semicircle MAM′ be such as to extend over two wave-lengths of the incident light; the semicircle represents the plane (oblique) section through a wave surface lying in air; the plane of the semicircle is the plane of incidence, and CA is the normal at the point of incidence.

On entering the Iceland spar at C the ray of light is doubly refracted, giving rise to the ordinary ray CO (for which the refractive index is 1·65) and the extraordinary ray CE (for which the refractive index lies between 1·65 and 1·48). HBH′ is a semicircle drawn in the plane of incidence, GBG′ an ellipse; the two together form the section of the Fresnel wave surface by the plane of incidence. Of the two components the ordinary ray is polarized in the plane of incidence, i.e. vibrates in a plane at right angles to the plane of incidence, as shown by the arrow at O, while the extraordinary ray is polarized at right angles to the plane of incidence, i.e. its plane of vibration coincides with the plane of incidence, as indicated by the arrow at E. The magnitude of the two curves representing the wave surface is again so chosen that the distance from centre to circumference represents two wave-lengths.

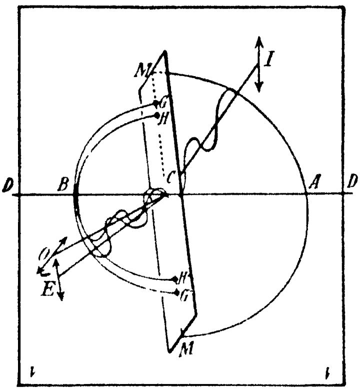


Fig. 35.—Model to explain the interference figures produced by plates of uniaxial crystals between polarizer and analyser

The amplitudes of the two components depend on the angle made by the plane of vibration of the incident ray with the plane of incidence. We now imagine the whole central portion of the figure gradually rotated about the axis DD, in order that the plane of incidence may occupy all possible positions; the plane of vibration of the incident polarized ray, however, is always to remain *vertical*.

(1) If the plane of incidence is vertical, it coincides with the plane of vibration of the incident ray; hence the extraordinary ray alone enters the Iceland spar, while the ordinary ray has zero intensity. The incident ray accordingly goes through the plate of Iceland spar entire as the extraordinary ray, vibrating with its full amplitude. For all rays polarized in a plane at right angles to the plane of incidence the plate of Iceland spar accordingly behaves like a singly-refracting medium.

(2) If the plane of incidence is horizontal, it is at right angles to the plane of vibration of the incident ray; hence the ordinary ray alone enters the Iceland spar, and with its full intensity, while the extraordinary ray has zero intensity. Hence for all rays polarized in a plane parallel to the plane of incidence the plate of Iceland spar again behaves like a singly-refracting medium.

Hence as there is no splitting-up into two components in the vertical and horizontal planes of incidence, in these planes the field of view is bright when the polarizer and the analyser are parallel (giving the bright cross in fig. 33) and dark when the polarizer and analyser are crossed (giving the dark cross in fig. 34).

(3) For every other plane of incidence the incident ray is broken up into two components; according to the angle which they form with the axis, they acquire a different phase difference on passing through the plate of Iceland spar. When their planes of vibration are made the same again by means of the analyser, they interfere with one another as the components did in the gypsum plate. Hence in every direction a definite colour is produced. All the rays which form the same angle with the axis, however, exhibit the same colour; since their components have the same phase difference, the colours are arranged in rings. The colours are deepest for those planes of incidence which form an angle of 45° with the planes of the polarizer and analyser, since for these planes of incidence the two components $a \sin \varphi$ and $a \cos \varphi$ have the same amplitude

$a\sqrt{2}/2$, while the amplitudes differ the more the closer the planes of incidence approach the vertical or horizontal plane; in these directions the colours accordingly become steadily paler.

The whole phenomenon is referred to as the *rings and brushes* produced by uniaxial crystals. Biaxial crystals give rise to other figures of a similar nature, but we shall not give details here.

Double Refraction of Isotropic Media arising from Strains.—Double refraction arises from the variation of the rate of propagation of light in different directions; this depends on the molecular structure of the substance, so that double refraction is always produced when the normal isotropic structure of a body is altered by external conditions. Glass plates become doubly refracting if they are subjected to pressure on one side or cooled suddenly and irregularly after heating. If stressed or rapidly cooled pieces of glass are placed between a polarizer and an analyser, very curious figures are often seen, from the shape of which the internal stresses may be inferred (fig. 36, Plate XVI). Before being ground to form lenses and prisms optical glass is tested by a polariscope; if it is to be used for good quality instruments, it must not show any coloured bands or curves. Conversely, attempts have been made to gain information about internal stresses by making a glass or celluloid model of the body under investigation and examining it in polarized light under the conditions of stress to which the body is to be subjected (fig. 37, Plate XVI).

Investigation of Rocks and Minerals.—The polariscope is also used to investigate minerals and rocks. For this purpose a thin section of the substance is placed between the polarizer and the analyser and examined by parallel or convergent light. From the nature of the double refraction it is often possible to draw accurate conclusions about the nature of the mineral or composition of the rock.

The Weigert Effect.—If ordinary photographic paper which has previously been allowed to darken in ordinary light is intensely illuminated by polarized light of a single colour (e.g. red), part of the paper takes on the colour in question; the light reflected from this region is polarized, as is found on examination by an analyser, the brightness being a maximum when the plane of polarization is the same as that of the incident light, and a minimum when the two planes are at right angles. Other photosensitive layers also exhibit this effect, so that in a certain sense it is possible to photograph the plane of polarization (Weigert, 1919).

Haidinger's Brushes.—With some practice polarized light may be detected as such by the eye directly. If one looks through a polarizer at a perfectly uniformly illuminated surface, a peculiar figure, something like a dark yellowish circle with two bright bluish regions inside, is formed within the eye for a few moments (the so-called *Haidinger's brushes* *).

7. Rotation of the Plane of Polarization (Optical Activity).

The Rotation of the Plane of Polarization in Quartz.—Quartz, which crystallizes in the hexagonal system, behaves in a peculiar way in polarized light. If a quartz plate cut at right angles to the optic axis, i.e. at right angles to the principal crystallographic axis, is placed between the crossed polarizer and analyser of a polariscope, the field of view, which was previously dark, becomes bright; if white light

* See p. 185.

is used the quartz plate appears coloured. If the analyser is now rotated, the colour of the quartz plate varies and at the same time the intensity of the light diminishes, without, however, reverting to zero.

If *monochromatic* light is used there is merely an alternation of light and darkness. If a quartz plate is brought between a crossed polarizer and analyser, the field of view, originally dark, becomes bright; by suitable rotation of the analyser, however, complete darkness may again be obtained. The phenomena is explained by the assumption that *the plane of polarization of the light is altered* as the light passes through the quartz. As may be definitely established with thin quartz plates (up to about 3 mm. in thickness), there exist two kinds of quartz, which are mirror images of one another even in their crystalline form, and rotate the plane of polarization of light in opposite directions. If the quartz rotates the plane of polarization to the right, it is said to be " right-handed "; if it rotates the plane of polarization to the left, it is said to be " left-handed ". Most of the quartz found in nature in the form of rock crystal is *right-handed*, whereas smoky quartz is usually *left-handed*.

In monochromatic light the rotation required to extinguish the light varies with the wave-length. To produce darkness again, using a quartz plate 1 mm. thick, the analyser must be turned through an angle of 15° for red light, 21° for yellow, 27° for green, 33° for blue, and 51° for violet. The rotation of the plane of polarization in quartz accordingly differs for light of different colours, being least for the rays of longer wave-length and greatest for rays of shorter wave-length. The rotation depends also on the thickness of the plate of quartz. A plate 2 mm. thick rotates the plane of polarization through twice the angle that a plate 1 mm. thick does.

The rotation of the plane of polarization is illustrated diagrammatically in fig. 38. Let the light from the source L be polarized by the polarizer P in such a way that it vibrates in the direction of the arrow *b* (that is, the plane of polarization is the plane of the figure). Let the analyser A and the polarizer P be crossed. That is, let the analyser transmit only light vibrating in the direction of the arrow *a*. To the observer the field of view then appears dark. If a left-handed plate of quartz about 3 mm. thick is inserted at *xx* and red light is used, the analyser must be rotated until the arrow attached to it reaches the position *r* (through an angle of 45°) to make the field dark again. If violet light is used, the analyser must be rotated until the arrow reaches the position *v* (through an angle of 153°).

The colouring of the quartz plate in white light follows immediately from the variation of the rotation with the wave-length. If, for example, we insert a quartz plate 1 mm. thick between crossed polarizer and analyser and rotate the analyser through 15°, the red light is completely extinguished, while the other parts of the light are more or less enfeebled in transmission. The field of view will therefore appear greenish-blue. If the analyser is rotated through 21°,

the yellow light is extinguished and the plate appears blue. If the analyser is rotated through 27°, the green light is extinguished and the field of view appears purplish-red, and so on. The quartz plate accordingly exhibits the following succession of colours: green, blue, red, orange, yellow. If the analyser is rotated still further, the colours recur in the same succession.

If we use a quartz plate several millimetres thick, the colour of the plate appears to be less deep, as various parts of the spectrum then overlap and give rise to dull mixtures of colours, such as arise in all interference phenomena when the layer producing interference is made too thick.

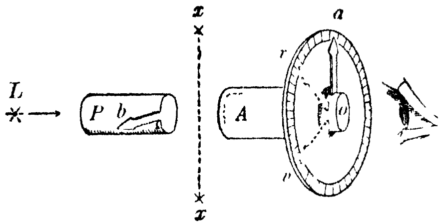


Fig. 38.—Rotation of the plane of polarization

The angle through which the plane of polarization is rotated by a plate of a solid substance 1 mm. thick is called the **rotation** (α). The variation of the rotation for the various colours of the spectrum is referred to as **rotatory dispersion** (see the table at the end of the book, p. 282).

Besides quartz there are a large number of solid bodies, as well as pure liquids, which rotate the plane of polarization. The rotation is particularly large in the case of liquid crystals (Vol. I, p. 335), where the value of α for the D line may reach 17,000°. Solutions may also possess the power of rotating the plane of polarization. For further details see § 9, p. 254.

Rotation of the Plane of Polarization by Sugar Solution.—If a glass tube bounded at either end by parallel planes is filled with sugar solution and the tube is placed in the path of the polarized light so that the latter passes through the two end surfaces, the plane of polarization is rotated. The amount of the rotation is proportional to the length of the tube and the concentration of the sugar solution.

The rotation of the plane of polarization by a sugar solution may be demonstrated very simply by leading a parallel pencil of polarized light through a glass tube about 1 m. long filled with a concentrated solution of sugar which has been made turbid by the addition of a minute trace of mastic solution. Here each particle of mastic acts as an analyser, like the glass pyramid in fig. 11 (p. 226); as a result the glass tube appears to be traversed by a bright spiral line. If a thin plate of quartz is inserted where the light enters the glass tube, the spiral is differently coloured in different azimuths.

The power which an optically active liquid or an optically active substance dissolved in an inactive liquid has of rotating the plane of polarization is measured by the **specific rotation** [α], which is defined as follows: If q is the number of grammes of the substance in 100 c.c.

of solution, l the length of the tube of liquid in decimetres, and α the observed angle of rotation, we have

$$[\alpha] = \frac{100\alpha}{lq}.$$

The specific rotation of cane sugar dissolved in water for sodium light ($[\alpha]_D$) is 66·5; that is, the angle (α_D) through which the plane of polarization is rotated by a column of liquid l dm. long containing z gm. of cane sugar per 100 c.c., is given by

$$\alpha_D = 0{\cdot}665^\circ \times l \times z.$$

Here again the rotation is different for different parts of the spectrum.

The rotation caused by a given length (l) of sugar solution may be used to measure the concentration of the solution, i.e. the number of grammes of sugar (z) contained in 100 c.c. of the solution. From the above formula it follows that for sodium light

$$z = 1{\cdot}503\,\frac{\alpha_D}{l}.$$

The Polarimeter.—This fact is made use of in the construction of instruments known as *polarimeters* or *saccharimeters*. These consist essentially of a source of light, a polarizer, and an analyser provided with a graduated circle, the last two being so arranged that a tube containing sugar solution may be inserted between them.

Fig. 32 is a diagrammatic sketch of a polarimeter due to MITSCHERLICH.* In this instrument accurate setting of the analyser is very uncertain, as it is difficult to judge when the field of view is completely dark. For this reason various additional devices have been introduced to enable a sharp adjustment to be made. Of these we shall mention the following:

Soleil's † **biquartz.** This consists of a right-handed plate of quartz and a left-handed plate of quartz 3·75 mm. thick, cemented together so as to exhibit a sharp line of demarcation. A plate of quartz 3·75 mm. thick rotates yellowish-green light through about 90°; monochromatic sodium light is rotated through 81·5°. Hence if the biquartz is inserted between parallel Nicols as in fig. 40, the yellowish-green light is completely extinguished by the two plates, and if the source of light is yellowish-green the field of view becomes completely dark. If white light is used, the whole field of view appears to have the complementary colour, reddish-violet. If the analyser is now rotated through a very small angle, the field of view becomes bright if the source of light is yellowish-green; if white light is used, one half of the field of view appears distinctly reddish, the other definitely blue. This difference in colour between the two halves of the field of view is very readily recognized, so that it is easy to adjust the analyser accurately to the position where both halves of the field of view have the same

* E. MITSCHERLICH (1794–1863), a noted chemist.

† N. SOLEIL (1798–1878), a French physicist.

reddish-violet colour ("sensitive tint"). Herein lies the importance of Soleil's biquartz when white light is used. If the light is monochromatic no advantage is gained by its use.

In Soleil's **saccharimeter** a similar biquartz is used, the sugar solution to be investigated being also placed between the polarizer and the analyser.

Instead of rotating the analyser, we use **a quartz wedge compensator.** This consists of a plate of right-handed quartz bounded by parallel planes and two

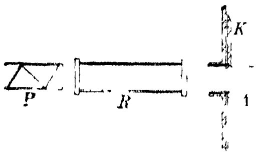


Fig. 39.—Diagrammatic sketch of Mitscherlich's polarimeter. P, polarizer; R, tube with solution; A, analyser; K, graduated circle.

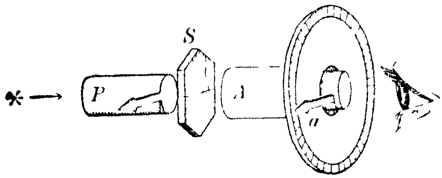


Fig. 40.—Polarimeter with biquartz

exactly similar wedge-shaped plates of left-handed quartz placed facing one another. If the two wedges of left-handed quartz are superposed, they act like a plate of left-handed quartz bounded by parallel planes. If the wedges are displaced relative to one another, the thickness of the plate is altered. We can then arrange that this double wedge of left-handed quartz has the same thickness as the right-handed quartz. In this case their effects cancel completely. If, however, one of the wedges is displaced in one direction or the other relative to the other, the plane of polarization is rotated in one direction or the other.

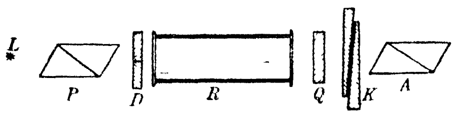


Fig. 41.—Polarimeter with quartz wedge compensator

By suitable adjustment of the quartz wedges it is accordingly possible to compensate for a rotation of the plane of polarization produced in any manner.

Fig. 41 shows the essential parts of a Soleil saccharimeter; in this figure P denotes the polarizer, D the biquartz, R the tube with sugar solution, Q the flat plate of right-handed quartz, K the double wedge of left-handed quartz, and A the analyser. The apparatus also contains a colour regulator and a small observation telescope not shown in the figure.

In some other types of optical saccharimeter the sensitiveness of the adjustment is magnified to an extraordinary extent, but we cannot go into details of these here.

8. The Production of Elliptically Polarized Light and Circularly Polarized Light.

When a pencil of polarized light passes through a plate of gypsum as illustrated in fig. 29 (p. 242), it will subsequently pass through the analyser, although the analyser and the polarizer are crossed. This depends on the fact that one component executes half a vibration less in the gypsum than the other does. If the analyser is not present, the ray split up by the gypsum emerges again

as a single ray arising from the combination of two light waves vibrating in planes at right angles to one another.

The resultant electric or magnetic intensity of the wave motion resulting from the combination of the two vibrations is formed from the two components of the ordinary ray and the extraordinary ray by the general method applicable to all vibrations, which is discussed on p. 195 of Vol. II. Here we are concerned with the special case where the two components have the same period. Further, we shall assume that we have to deal with the simple case where the amplitudes of the two vibrations are equal. The combination of the vibrations is carried out as shown in fig. 11 (Vol. II, p. 196).

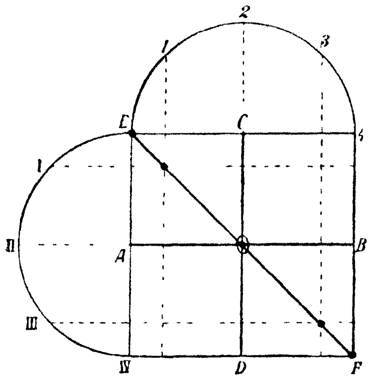


Fig. 42

In fig. 42 AB and CD represent twice the amplitude ($2a$) of the component vibrations about the centre of vibration O; the vibrations are reversed at the same instant at A and C respectively.

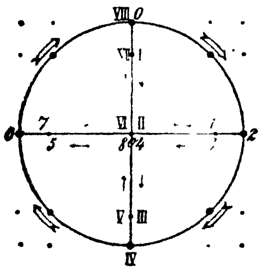


Fig. 43

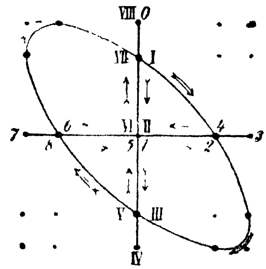


Fig. 44

Combination of linearly polarized light to form linearly polarized light, circularly polarized light, and elliptically polarized light

The resultant vibration is the linear vibration EF, whose azimuth is rotated through an angle of 45° relative to that of the components, and whose amplitude is $a\sqrt{2}$.

If the vibration AB begins at A at the instant when the vibration CD begins at D, the resultant vibration is also a linear vibration, but along the diagonal IV4. The amplitude is again $a\sqrt{2}$, but the azimuth makes an angle of 90° with that of EF.

Fig. 43 shows the case where the one vibration (AB in fig. 42) is passing through the central point O at the instant when the other vibration (CD in fig. 42) is starting from C. The difference of phase is $\pi/2$. In fig. 43 corresponding points of the two vibrations are indicated by corresponding numbers 0, 0; 1, I. 2, II, &c. The resultant vibration is a circular motion in which the rotation is right-handed (clockwise when looked down on). If one of the components has its phase altered by half a vibration, the circular motion is left-handed (counter-clockwise). (See also the analogous remarks about a rotating magnetic field on p. 503 of Vol. III.)

In fig. 44 the difference of phase between the two components is $\pi/4$. When

combined they give rise, as is the case in general when the phase difference is not exactly 0 or $\pi/2$, to an elliptical vibration.

An elliptical vibration is accordingly the most general form of resultant vibration.

If we apply these results to the combination of two coherent linearly polarized rays of light, vibrating in planes at right angles to one another, we see that their resultant is a *linearly polarized* ray, a *circularly polarized* ray, or an *elliptically polarized* ray, according to the difference of phase of the two components.

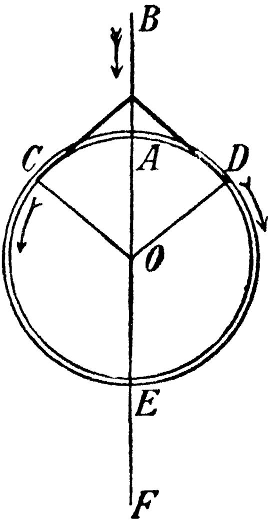


Fig. 45.—Combination of two circularly polarized vibrations to form a linear vibration

Linearly polarized light is always produced when the difference of phase of the two linearly polarized components is either zero or an even multiple of a quarter vibration, i.e. $2n\pi/2$. Circularly polarized light is produced when the difference of phase of the two components is an odd multiple of a quarter vibration.

It is easy to understand why the linearly polarized light again gives rise to linearly polarized light whose plane of polarization is rotated through 90° from that of the incident light, in the case shown in the model of fig. 29, p. 212; for one component is displaced relative to the other by half a wave-length, i.e. is displaced in time by half a period. If the displacement is a quarter of a wave-length or an odd multiple thereof, the two emerging components each of which is executing linear vibrations (is linearly polarized), must give rise to a circular vibration (circularly polarized light). This may be attained in practice by means of a plate of gypsum of suitable thickness.

A sheet of gypsum which gives the components a difference of phase of a quarter vibration is called a **quarter-wave plate.** If a quarter-wave plate is placed in the path of a linearly polarized pencil of light in such a way that the principal direction of vibration in the gypsum is inclined at an angle of 45° to the direction of vibration of the polarized light, the linearly polarized light is transformed into circularly polarized light. If two quarter-wave plates are superposed, the second transforms the circularly polarized light back into linearly polarized light vibrating in a plane at right angles to the plane of vibration of the incident pencil of light. A third quarter-wave plate will transform this light back into circularly polarized light, but now the sense of rotation of the circular motion is opposite to that of the circularly polarized light produced by a single quarter-wave plate.

Combination of Two Circular Vibrations.—Just as we may combine two components linearly polarized in directions at right angles to one another to produce circularly polarized light, we may also recombine two circularly polarized components. If the two components have the same amplitude and opposite senses of rotation, we get back linearly polarized light (fig. 45). The azimuth (BF) of the linearly polarized light thus formed, i.e. the position of the plane of vibration, depends on the points at which the two circular vibrations meet.

If the two circular vibrations have the same period, they always meet at the two extremities of the same diameter; this diameter fixes the azimuth of the linear vibration.

If the period of one of the circular components, say that with right-handed rotation, is somewhat less than that of the other, i.e. its velocity is somewhat

greater than that of the other, the azimuth will move through a definite small angle at each revolution in the direction of the more rapidly vibrating component. Hence the plane of polarization of the linearly polarized light gradually rotates. This case does not arise directly in nature, as a difference in period between the two components would entail a difference in their colours, although the components would still have to be coherent.

Such a difference may arise from the effect of strong magnetic fields within a material medium traversed by light; see also p. 256.

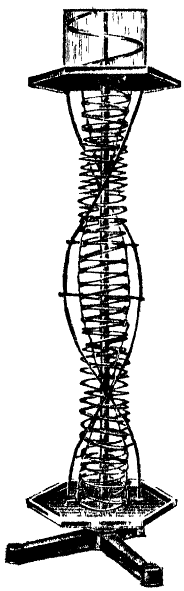


Fig. 46.—Model to explain the rotation of the plane of polarization

9. Explanation of the Rotation of the Plane of Polarization.

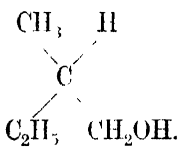
Following FRESNEL, we may explain the rotation of the plane of polarization as follows:

If a linearly polarized pencil of light traverses a quartz plate, cut at right angles to the optic axis, in the direction of the optic axis, the linearly polarized light is split up into two circularly polarized rays of light with the same amplitude and opposite senses of rotation. The right-handed component, however, is assumed to advance in the direction of the light ray more slowly than the left-handed one.

This may be illustrated by means of the model shown in fig. 46. On a wooden cylinder we wind two similar wire spirals one on top of the other, one being right-handed, the other left-handed. The two spirals have the same number of turns per unit length. One end of each spiral is fixed to the wooden cylinder, and we pull the other ends, stretching the left-handed spiral so that it is somewhat longer than the right-handed one. Then the right-handed spiral represents the right handed circularly polarized component of the light and the left-handed spiral the more rapidly advancing left-handed circularly polarized component. The points where the two superposed spirals touch determine the azimuth of the linearly polarized light arising from the two components at any particular instant. We see that the azimuth rotates in the direction of a steep right-handed spiral. This explanation is in good agreement with the observed behaviour of quartz. For example, if on one side of a quartz crystal we grind a face inclined at a very small angle to the optic axis, and on the other side a face normal to the optic axis, a linearly polarized ray of light incident on this face in the direction of the optic axis must enter the crystal without refraction. Within the quartz the ray then breaks up into a right-handed circularly polarized ray and a left-handed circularly polarized ray, the wave velocities in the two rays being different. As the two rays no longer fall normally on the polished face first mentioned, they are differently refracted as they emerge, that is, there emerge from the surface inclined to the optic axis *two* circularly polarized rays inclined to one another instead of a single linearly polarized ray. Thus quartz exhibits double refraction—although to a very trifling extent—even in the direction of the optic axis.

The Structure of Optically Active Substances.—The rotation of the plane of polarization is to be ascribed to the fact that the structure of the elements of the crystal or even of the molecule (e.g. in liquids

and solutions) exhibits a spiral structure so far as the resonators (p. 171) on which the propagation of light in the substance depends are concerned. There exist dextro-rotatory and lævo-rotatory substances, according to the sense of rotation of the spirals. Substances of the same chemical constitution sometimes occur in three different forms, a dextro-rotatory form (the so-called *d* form), a lævo-rotatory form (*l* form), and an inactive form (racemic form). The occurrence of this so-called **optical activity** in the case of organic compounds is associated with the presence of an asymmetrical carbon atom, i.e. a carbon atom which has a different group of atoms attached to each of its four valencies, as is the case e.g. in amyl alcohol,



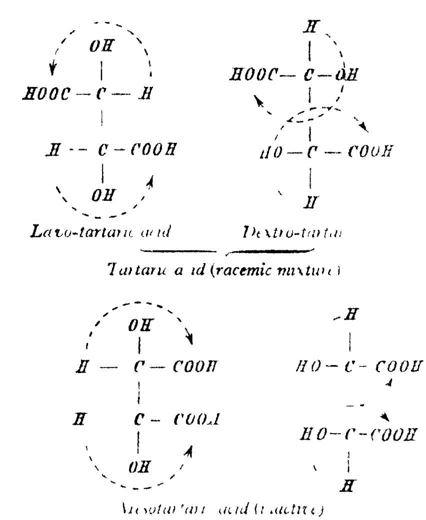


Fig. 47.—Projection on a plane of the groups of atoms in tartaric acid regarded as being situated at the vertices of a tetrahedron

If the central atom is thought of as being situated at the centre of a tetrahedron, the above groups may be distributed among the vertices of the tetrahedron in two different ways which are not superposable. The racemic form consists of a mixture of the two components in equal quantities, so that their rotations exactly cancel. In the case of tartaric acid, the molecule of which contains *two* asymmetrical carbon atoms (two active groups), compensation may also arise owing to the senses of rotation *within* the molecule being opposite. This form (mesotartaric acid) is actually known (fig. 47).

The correctness of these ideas has been proved by experiments with metal spirals. A large number of identical spirals with the same sense of rotation were placed in a box, forming a model of an optically active substance. When electric waves of suitable wave-length were made to traverse the box, a rotation of their plane of polarization was detected and found to agree even quantitatively with the calculated amount.

10. The Effect of Electric and Magnetic Fields on the Propagation of Light.

The Effect in a Vacuum.—In a vacuum even the strongest electric and magnetic fields are entirely devoid of any effect on the propagation of light. This shows that electric and also magnetic intensities and displacements are superposed in space free of matter without the least mutual disturbance. A particular case of this has already been mentioned on p. 9.

The Effect in Space filled with Matter.—The light is affected in the following way: under certain circumstances even the atoms emitting the light are set in certain definite directions, so that the "ordinary" light starting from the source is no longer unpolarized, but exhibits polarization in a definite direction owing to the elementary oscillators taking up certain favoured positions relative to the external field.

Light once emitted is liable to be affected during its passage through matter by the presence of an electric or magnetic field.

Effects taking place on emission.

I. *If the source of light is brought into a strong magnetic field, the emitted light* (1) *has its wave-length altered from its value in the absence of a field,* (2) *is polarized* (the **Zeeman** * **effect** (1896)).

II. *If light is emitted in a strong electric field, the light* (1) *has its wave-length altered,* (2) *is polarized* (the **Stark** † **effect** (1913)).

In both cases the direction of polarization and the change in wave-length are connected by a definite relationship.

These two phenomena are of fundamental significance in connexion with our ideas of the nature of the process of light emission. In view of their important bearing on our knowledge of the structure of the atom they will be discussed in detail in Vol. V.

Effects taking place when Light is moving through Space containing Matter.—In the following cases the effect of external fields is to be thought of as due to the elementary resonators (p. 171) being individually affected by the field according to the Zeeman or Stark effect, so that the elementary waves starting from them have their wave-length and direction of polarization changed. An explanation of the phenomena described below will be given in Vol. V.

* PIETER ZEEMAN, born in 1866, was engaged in physical research at the University of Amsterdam. The discovery mentioned above was made by him at Leyden in 1896 as the result of a suggestion by H. A. LORENTZ. In 1903 the two investigators were awarded the Nobel prize for physics. FARADAY had already sought for an effect of the magnet on the colour of light, but was unsuccessful owing to the fields he used not being intense enough. Observations similar to those of ZEEMAN had been made by C. FIEVIEZ as early as 1885, but they had been lost sight of.

† JOHANNES STARK, till 1917 Professor of Physics at Aix-la-Chapelle, at Greifswald from 1917 to 1920, and at Wurzburg from 1920 to 1921, was awarded the Nobel prize in 1919, and became president of the *Physikalisch-technische Reichsanstalt.*

Electric Fields. *Electrical Double Refraction* (the electro-optic **Kerr** * **effect**).—If light is made to pass between the plates of a flat condenser filled with a suitable substance, e.g. nitro-benzene, this substance will exhibit double refraction when an electric pressure is applied to the condenser. Two ray-components are produced, one being polarized parallel to the electric field, the other at right angles to the electric field. The components are propagated at different rates. A light ray polarized in a direction making an angle of 45° with the lines of force is therefore elliptically polarized after traversing the liquid (p. 253). Hence if an analyser is adjusted so that the ray of light traversing the so-called Kerr cell is extinguished when the electric field is absent, the field of view will become bright again when the electric field is applied. The difference of path between the two rays in wave-lengths is

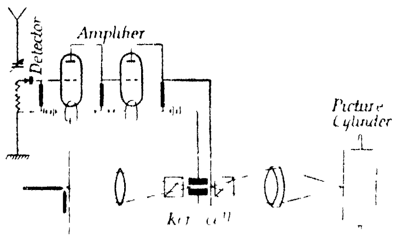


Fig. 48.—Principal circuit of a receiver for wireless transmission of pictures

$$k = KlE^2,$$

where K is the so-called Kerr constant of the substance used, l the thickness of the layer of substance in centimetres, and E the electric intensity in volts/cm

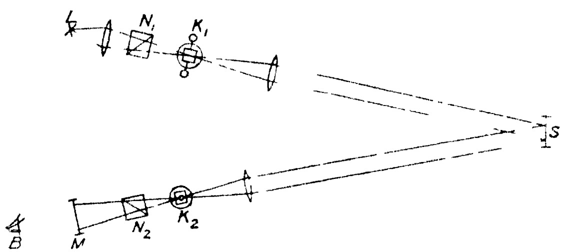


Fig. 50.—Principle of the method for measuring the velocity of light by means of the Kerr cell: L the source of light, K_1 and K_2 Kerr cells, B the observer, N_1 and N_2 the polarizers, S a mirror, M a disc of matt glass (from *Annalen der Physik*, Vol. 2 (J. A. Barth, Leipzig)).

That is, the brightness of the field of view depends on the intensity of the electric field. We may also cause darkness when the polarizer and analyser are parallel by switching on the field. The Kerr cell has become of great practical importance in connexion with the wireless transmission of pictures and with television, largely as a result of the work of Karolus.† It enables variations in an electric field to be transformed without any lag into variations of light, so that the principle underlying its application to the above purposes is obvious; the variations in the amplitude of the oncoming electric waves (i.e. the variations of the electric field) are transformed into variations of light, which are recorded on photo-sensitive paper as in fig. 48. The substance chiefly used is nitrobenzene, which has a Kerr constant of $2{\cdot}7 \times 10^{-8}$ for the D line. It was with glass that Kerr discovered the phenomenon.

Determination of the Velocity of Light by means of the Kerr Cell (Karolus

* Discovered in 1875 by John Kerr (1824–1907), a Scottish physicist.

† Professor at the University of Leipzig.

and MITTELSTAEDL, 1928). If an alternating pressure, or better still an alternating pressure superposed on a direct pressure (Vol. III, p. 561), is applied to a Kerr cell, and the values of E and l are suitably chosen, the light traversing the cell is interrupted with a frequency double that of the alternating current (as it depends on E^2). Hence, as was suggested by DES COUDRES as early as 1893, a Kerr cell could be used in Fizeau's experiment instead of the toothed wheel. With modern methods of producing vibrations it is possible to hold a frequency of about 10^7 hertz (cycles) steady to within $\pm$ 200 hertz. This enables the path of the light to be cut down to less than 100 m. The measurements are made by a compensation method (fig. 50). If in the path of a linearly polarized pencil of light we place two crossed cells (i.e. at an angle of 90° with each other), we may, as considerations like those on p. 253 show, cancel the double refraction produced in the first cell, provided that the two cells are exactly the same, and have the lines of force inclined at an angle of 45° to the plane of polarization. With crossed cells, therefore, we obtain darkness in every case. If, however, the light takes a considerable time to pass from one cell to the other, the applied pressure in the second cell is in a different phase. By altering the frequency of the applied alternating pressure, it is possible to arrange for the second cell to have exactly the same phase as the first at every instant for the ray of light traversing the path s; the field of view is then dark again all the time. If the path s is kept constant the necessary frequency f is accordingly given by $f = c/s$, or more generally by $f = kc/s$, where k is a whole number. With a path s of about 300 m. (obtained by repeated reflection, k taking values from 4 to 8), the average value

$$299{,}778 \text{ Km./sec.} \pm 20 \text{ Km./sec.}$$

was obtained for the velocity of light.

Magnetic Fields.—*Magnetic Rotation of the Plane of Polarization.*—In 1845 FARADAY discovered the first-known relationship between magnetic (and hence also electric) phenomena and optical phenomena. He found that the plane of polarization of linearly polarized light is rotated in a magnetic field in the presence of matter. To carry out Faraday's experiment a piece of glass consisting largely of lead silicate (Faraday's glass) and having its end surfaces ground flat and polished is placed between the poles of an electromagnet so as to lie along the lines of force. Nicol prisms are placed at either end of a hole bored through the pole-pieces of the magnet to act as polarizer and analyser; the two Nicols are crossed, so that to an observer looking through the analyser the field of view appears dark. If the circuit of the electromagnet is then closed, the field of view lights up, but darkness may be reobtained by rotating the analyser, the angle through which the analyser is rotated giving the rotation of the plane of polarization. The sense of rotation is that in which the electric current flows round the electromagnet. A rotation in this direction is reckoned positive. Rotation of the plane of polarization is also found to occur in other substances; in some the rotation is negative. The magnitude of the rotation (α) is proportional to the intensity of the field as well as to the length of substance traversed by the light:

$$\alpha = \omega l H,$$

where H is measured in gauss. Further, the rotation generally increases as the wave-length diminishes.

The constant ω means the angle through which the plane of polarization of light is rotated when the length of substance traversed is 1 cm. and the intensity of the magnetic field is 1 gauss. This quantity is known as **Verdet's * constant.**

* M. E. VERDET (1824–66), a French physicist particularly distinguished by his researches in optics.

If we measure the rotation in seconds of arc, Verdet's constant is 3·4 for Faraday's glass, 4·4 for molten sulphur, 7·3 for phosphorus, 2·5 for carbon disulphide. An aqueous solution of ferric chloride gives a negative rotation; Verdet's constant for a 40 per cent solution of ferric chloride is −5. KUNDT * (1884) investigated extremely thin metallic films (so thin as to be transparent) in the magnetic field, and found that metals also rotate the plane of polarization in the magnetic field. The ferromagnetic metals iron, nickel, and cobalt give a particularly large rotation; the rotation of iron is about 30,000 times that of glass. Further, it is noteworthy that in the case of iron the rotation is greater for red light than for blue light; that is, there is anomalous rotatory dispersion.

KERR observed in 1877 that the plane of polarized light is also rotated when it falls on the polished end surface of a strong magnet and is reflected from it (the magneto-optic **Kerr effect**).

Magnetic Double Refraction.—This is the magnetic analogue of the electric double refraction discussed on p. 257. As in that case, we have

$$k = ClB,$$

where C is the so-called ***Cotton-Mouton constant*** of the substance concerned, l the thickness of the layer traversed by the light, and B the magnetic induction (measured in gauss). The value of C for nitrobenzene, in which this phenomenon is well developed, is $2·41 \times 10^{-12}$ for the D line.

* AUGUST KUNDT (1838–1914), Professor of Physics at Bonn.

CHAPTER XI

Optical Phenomena in the Atmosphere

1. Atmospheric Refraction.

When a light ray passes from a vacuum to air it is refracted towards the normal. At 0° C. and 760 mm. pressure the refractive index of air is 1·000291. A rise in temperature or a fall in pressure diminishes the density and hence also the refractive index of air. If the normal density of air is d and the normal refractive index μ, the refractive index μ' when the density is d' is given by the equation

$$\frac{\mu' - 1}{\mu - 1} = \frac{d'}{d}.$$

A ray of light which travels over long distances in the terrestrial atmosphere must traverse various layers of air of increasing or decreasing density.

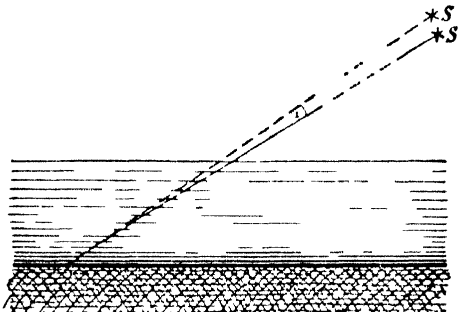


Fig. 1.—Atmospheric refraction

Each time it passes from one layer to another it is subject to refraction. Hence its path through the atmosphere is a slightly curved line (p. 107). All rays of light in air, therefore, are really slightly curved. If the distance is small, however, this curvature cannot be detected; if the distance is considerable, it is one of the sources of error well known to the surveyor, which must be taken account of in accurate measurements. Rays of light reaching the earth from a star are likewise slightly curved (fig. 1). As by Chap. I, § 2,

p. 4, the eye infers that the source of light is situated on the backward prolongation of the ray of light entering the eye, the star appears higher to the observer than it really is (S′ in the figure indicating the apparent position, S the real position). The angle (α) through which the source is apparently raised, measured in radians, is called the **atmospheric refraction.** It is greater the lower the star is in the sky. The atmospheric refraction of a star on the *horizon* is called the **horizontal refraction**; this amounts to about 0·5°. The atmospheric refraction for various altitudes is tabulated in the appendix (p. 282).

As a result of atmospheric refraction we are still able to see stars which in reality have already sunk beneath the horizon. At the instant when the *lower* rim of the setting sun is apparently touching the horizon, the *upper* rim has actually already sunk beneath the horizon. The same is true of the sunrise.

The flattened appearance of the sun or moon at rising or setting is largely to be ascribed to the fact that the light from the lower rim is more strongly refracted by the atmosphere than that from the upper rim.

An apparent raising of objects on the surface of the earth, analogous to the phenomenon just mentioned, is often observed on the banks of large lakes or at the sea-coast, and also on the open sea, when the layer of air next to the water is markedly cooler, i.e. markedly denser, than the layers above it. The rays rising obliquely from an object in the neighbourhood of the surface of the water are refracted away from the normal and may assume an almost horizontal direction. As a result rays may reach the eye from objects which actually lie below the horizon. For similar curvatures in the case of sound rays see Vol. II, fig. 1, p. 277, and in the case of electric waves Vol. III, fig. 45, p. 640.

As the refractive index of air depends on the density and is less for high temperatures than for low, the direction of a ray will also be altered by hot ascending currents of air, e.g. from a chimney or a wall or piece of ground exposed to the sun. The shape and position of these ascending currents of air are liable to vary irregularly. Hence objects looked at across these ascending currents of air appear to tremble or flicker. The flickering or **twinkling** of the fixed stars is also due to the variations in refraction of the light rays due to the variable density of the atmosphere. The wave-fronts for the light from the stars, which ideally should be plane, exhibit certain irregularities whose average size is about 100 sq. cm. and radius of curvature several thousand metres. As each curvature of the medium gives rise to divergence or convergence of the rays, marked fluctuations in brightness result.

Mirages.—In certain circumstances it may happen that low-lying layers of air have a lower density than those lying above them. This occurs e.g. when the earth's surface is brightly illuminated by the sun in the absence of wind, so that the layers nearest the ground are strongly warmed. In this case a pencil

of rays inclined at a very small angle to the ground is deviated away from the normal. We then have the phenomenon illustrated by the experiment of fig. 28, p. 107; the ray of light is curved. If the curvature is so marked that the ray reaches a lowest point and begins to travel upwards again, it appears to suffer total reflection and the hot layer of air acts like a glittering mirror. Owing to this apparent reflection, an observer can then see objects which lie above the ground twice, as they are apparently reflected in the mirror-like layer of air. Such a reflection may simulate the presence of an extensive sheet of water (**mirage, fata morgana** *). If the case is as shown in fig. 30, p. 108, two mirror images of the object are visible just below it, the upper being reversed relative to the object and the lower right side up. The two images are markedly compressed in the up and down direction. The formation of the two images and their distortion may be demonstrated by the experiment shown in fig. 2, in which, as in fig. 28, p. 107, the formation of a mirage is imitated by means of layers of alcohol and water.

A similar phenomenon may also be observed at the seaside on hot summer days. This occurs when a thin layer of water on the sea-bed (especially over parts exposed at low water) is strongly warmed; for this again gives rise to marked warming of the layer of air immediately above it. As the air at the same time is rarefied owing to the water vapour present, mirages often occur in still weather. We then have the impression that the level of the sea is raised; the apparent total reflection of the markedly rarefied layer of air gives rise to mirror images of objects on the land or on the surface of the water.

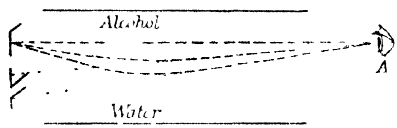


Fig. 2.—Imitation of the triple image of a mirage

Finally, at great altitudes an extensive layer of very rarefied air is frequently formed. Here total reflection following on refraction takes place in a similar way, and mirror images of objects on the ground are formed in the *sky*. In certain isolated cases it is thus possible to see objects which could not be seen directly at all owing to the curvature of the earth. Thus, for example, the mirror image of Heligoland has often been seen in the sky from Cuxhaven, although Heligoland itself cannot be seen from Cuxhaven.

In the **alpine glow** the rays of the sun are bent back after sunset by a very rarefied layer of air at a great height so as to illuminate the mountain tops directly.

2. Ordinary Diffuse Daylight.

The fact that we see not only those objects which are directly illuminated by the sun, but also objects lying in shadow, shows that in most cases the sunlight only reaches bodies after being deviated in some way or another. Sunlight is reflected, diffracted, and scattered by bodies on the surface of the earth, by the air (p. 232) and by bodies floating in it, such as water droplets, clouds, and particles of dust. The illumination of the earth is then uniform or **diffuse**, and does not give rise to such deep shadows as those cast by sunlight. In the open air the illumination produced by daylight is about 10,000

* *Fata morgana*, Arabic *fâmurgan*, a fairy who exhibits her power by causing mirages. The phenomenon is very frequently exhibited by slightly rising asphalt roads in sunshine.

lux (Chap. II, § 1, p. 23) i.e. is equal to an illumination of 10,000 candles at a distance of 1 m.

So long as the sun's rays can illuminate parts of the atmosphere which are visible to us, the sky is not completely dark. The period after sunset or before sunrise during which we receive diffused (scattered) light only is called **twilight.** From the duration of twilight it is possible to calculate to what height in the atmosphere particles of a size ($< \lambda$) capable of scattering sunlight extend. Twilight lasts longer the more acute the angle made by the path of the sun with the horizon; hence in high latitudes it is longer than in the neighbourhood of the Equator.

The Blue Colour of the Sky.—The light from the sun is scattered by the atmosphere; this is very probably due to the statistical fluctuations in the density of the air arising from the thermal motions of the molecules (Vol. II, p. 141). The blue colour of the sky is accordingly the colour of a "turbid medium" (p. 232). From the intensity of the scattered light Avogadro's number may be calculated; the value obtained by this method, $2 \cdot 89 \times 10^{19}$ molecules per cubic centimetre, is in good agreement with the results of other methods (Vol. II, p. 60).

As when the sun is low in the sky the light coming from it to our eyes has to traverse a very thick layer of air lying near the ground and hence containing a large number of very minute particles and water droplets, it is chiefly the red rays of long wave-length that get through, the other rays being eliminated during their long journey through the atmosphere owing to the scattering effect of small material particles, which increases in proportion to $1/\lambda^4$, by p. 232. The red rays of the rising or setting sun illuminate the parts of the sky near the horizon and produce the colours of sunrise and sunset, which are visible even when the sun is below the horizon.

If the particles floating in the atmosphere markedly exceed the size mentioned above, diffraction plays only a small part, and all the colours of sunlight are reflected and diffusedly scattered in apparently the same intensity; hence when small droplets have been formed by the condensation of water vapour and these have joined to form drops, the sky appears whitish. The white colour of mist or fog is also to be ascribed to the presence of small droplets of water. Similarly the smoke from a cigarette, which consists of very fine particles, is very distinctly blue, while the smoke exhaled, which consists of much coarser particles, is whitish.

The **polarization** of the light from the sky, which may be anticipated to occur in view of the discussion on p. 232, may be observed by pointing a Nicol prism at the sky and rotating the prism in front of the eye.

3. Solar and Lunar Coronæ and Halos.

Coloured Rings round the Sun or the Moon (Coronæ).—These are purely diffraction phenomena. Owing to the greater brightness of sunlight, coloured rings round the sun are less frequently observed than coloured rings round the moon. The moon appears surrounded by a bluish circular region bounded by a reddish-yellow circle. On

the outer side this may be followed by one or more successions of narrow coloured bands ranging from blue to red. While the first series of colours agrees more or less with the order of the colours in the spectrum, the succession of colours in the outer rings differs entirely from that in the spectrum. It follows from this that the phenomenon is not one of dispersion, but is due to the interference of light rays.

In appearance and mode of occurrence the phenomenon is very similar to the rings which appear to be formed round a distant source of light when it is looked at from a dark room through a slightly fogged window (e.g. one that has been breathed upon). Such rings are also formed when the piece of glass between the eye and the point source of light is slightly dusty, and are especially well shown when the dust is very uniform, i.e. consists of particles all about the same size (e.g. lycopodium powder, the spores of club-moss). If the glass plate is breathed upon the rings are at first large and brilliantly coloured; as the condensed droplets coalesce, the rings gradually shrink and the colours fade. The water droplets or particles of dust act like a diffraction grating. Just as the interference bands produced by the diffraction grating are farther apart the closer the ruling of the grating is, the coloured rings are larger the smaller the diffracting particles of dust or water droplets are. If the particles are of unequal size, the various diffraction rings overlap and destroy the purity of the colours.

This explains why coloured rings round the moon are chiefly observed on evenings after clear days; for then there is slight condensation of the water vapour present in the air, minute droplets of approximately equal size being formed. As condensation proceeds the water droplets become larger and unequal in size, and the ring becomes smaller and finally passes gradually into a whitish circular region surrounding the moon.

Solar and Lunar Halos.—These are essentially different from *coronæ*. Halos are always formed at the same distance from the sun or moon; usually they have a mean radius of 22°, more rarely of 46°; the inner side of the halo is always red, the outer side blue. Their occurrence is due to dispersion of the light at the small ice crystals formed in the higher layers of the atmosphere by the condensation of water vapour.

The ice crystals have the form of regular hexagonal columns with plane end faces. If a ray of light L (fig. 3) falls on one side face of a hexagonal ice crystal at an angle of 41°, it is refracted into the ice crystal at an angle of 30° (the refractive index of ice being 1·31) and subsequently leaves it again, the angles being interchanged. The ray is thus deviated through an angle of 22°. At the same time the ray is split up into its spectral components. Owing to the great distance between the crystal and the observer, however, only one ray of the fan of colours emerging from an ice crystal will enter the eye of the observer; hence any one crystal appears to send out only one definite colour. (This may readily be observed on walking over a sheet of snow on a cold sunny winter's day; here and there one sees a blue, green, or red flash from a snow crystal which happens to have the right situation relative to the sun and the eye.) As the blue rays are more strongly refracted than the red, the angular distance from the sun of those crystals which are apparently coloured blue is somewhat greater than that of the crystals which are apparently coloured red. For this reason the inner edge of the halo appears red and the outer edge blue.

The lunar halo of mean angular radius 46° is explained by the refraction and dispersion of light falling on the plane end surfaces of hexagonal needles of ice and emerging from their side faces. The path of the rays in this case is shown in fig. 4. The angle of incidence is 68°, the angle of refraction in the ice 45°, so that the deviation of the ray is 46°. Here again the deviation is somewhat greater for blue than for red, so that in this case also the inner edge of the halo is red and the outer edge blue.

Mock Suns.—In addition to the phenomena mentioned above, there often occur white rings due to *reflection* of light at the surfaces of the ice crystals. Occasionally these rings are accompanied by horizontal bands of light, which give rise to apparent expansions of the rings on either side of the sun or moon (*mock sun* or *moon*). These horizontal bands of light are due to the reflection of light at the plane

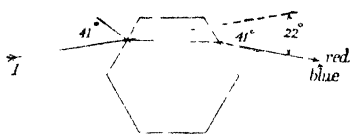


Fig. 3.—Deviation of light by a crystal of ice

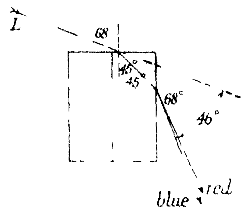


Fig. 4.—Deviation of light by a prismatic needle of ice

end surfaces of thin, flat, plate-like ice crystals, which take up a horizontal position in absolutely still air.

4. The Rainbow.

The Primary Rainbow.—A rainbow is formed when a cloud which is breaking up into *drops* is illuminated by the sun. The rainbow forms part of a circle whose centre is at the counter-sun, i.e. the point in the field of view which, as seen by the observer, is exactly opposite to the sun. The *primary rainbow* has on its outer side a fairly sharp red margin; inwards the colours succeed one another more or less in the order of the spectrum until blue is reached; after this there are usually a number of narrow coloured rings in which the succession of colours is apparently irregular (*supernumerary bows*). The *external* radius of the primary rainbow is always 42° 30′; the details of the succession of colours are different in almost every rainbow. The rainbow must depend partly on interference phenomena; for if it were a dispersion phenomenon pure and simple, only one succession of colours could occur and that always in the same order.

If we look at a very bright rainbow through a monochromatic red glass we see a succession of circular arcs, alternately bright and dark, similar to the diffraction rings which are formed when the light

from a point source (sunlight or a distant arc lamp) falls through a small circular stop on to a white screen.

The Secondary Rainbow.—This always has an *internal* radius of 51°; its inner edge is fairly sharply defined and is coloured red. Here the colours follow one another outwards in more or less the order in which they occur in the spectrum until blue is reached; then there are usually several supernumerary bows varying in number, breadth, and colour. Looked at through a glass transmitting one colour only, the secondary rainbow also appears as a system of bright and dark rings. In its whole appearance the phenomenon closely resembles the diffraction rings formed round the shadow of an opaque circular

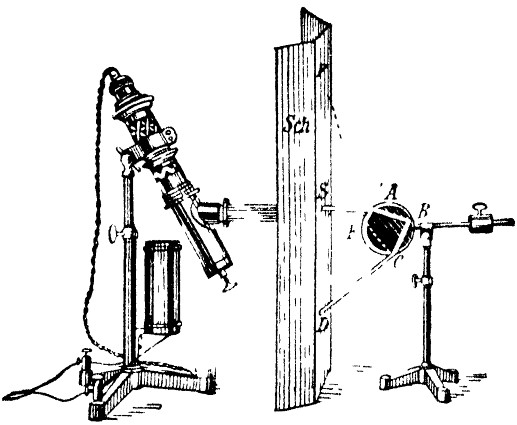


Fig. 5.—Grimsehl's apparatus for explaining the rainbow

disc placed in the path of the rays from a distant point source (fig. 18, Plate I).

Explanation.—To illustrate the phenomena of the rainbow we begin by investigating the course of a pencil of parallel rays falling on a sphere (or cylinder) full of water. A parallel pencil of (monochromatic) light is allowed to fall on the back of a white screen Sch provided with a narrow horizontal slit S (fig. 5), so that a narrow pencil is selected, which falls on the curved surface of a cylindrical vessel full of water. The cylinder may readily be raised or lowered so as to give the angle of incidence of the pencil of rays on the cylinder (which represents a drop of water) any value at will. We observe that the incident rays are partly reflected at the front surface of the drop (A) and partly refracted into the drop. The refracted portion of the light is again partly reflected at the rear side of the drop (B) and partly refracted out of the drop. The part of the light reflected at B is again partly reflected at C, the remainder of the light leaving the drop in the direction CD and meeting the screen at D.

If we let the ray SA fall on the drop at right angles, part of it is reflected into itself and part of it enters the drop without change of direction, some of the latter portion leaving the drop again in the same direction. If the drop is lowered, the ray CD is also lowered; CD makes an acute angle with AS, which if the drop is lowered farther at first rapidly increases, then increases more

slowly, and finally diminishes again. That is, the deviation of the pencil of rays depends on the angle of incidence, but never exceeds a certain large value. For red light it is 42° 30′; this value occurs when the angle of incidence of the light is 59° 30′. It follows that a pencil of rays (of red light) which illuminates the *whole* drop meets the screen only in a sharply defined zone whose boundary rays make an angle equal to the supplement of 42° 30′ with the original direction of the ray. Hence if parallel rays of (red) light fall on a single spherical drop, the rays which leave the drop backwards after being twice refracted and once reflected do so within a cone of semi-vertical angle 42° 30′. For light of shorter wave-lengths the semi-vertical angle of the cone is somewhat smaller (41° for blue light).

The cone of light in question is sharply bounded on the outer side, i.e. it behaves as if it were stopped down by a circular hole in an opaque screen. Hence at the edges we have the same diffraction phenomena as are produced by a screen with a sharp edge (fig. 6). Maxima and minima of intensity occur in alternate succession, the distances between them becoming smaller and the differences of intensity becoming less (cf. fig. 24, Plate I). This region gradually passes into a region of uniform illumination. In fig. 6 the brightness of the diffraction phenomena is plotted as ordinate; the point A corresponds to the boundary of the

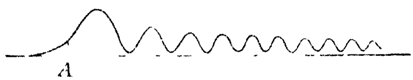


Fig. 6 — Maxima and minima of intensity in the rainbow when observed in monochromatic light

geometrical shadow, i.e. to the boundary of greatest deviation in the rainbow, that is, to a solid angle of 42° 30′ for red light and 41° for blue light. The distances between the various maxima of intensity are proportional to the wave-length of the light, and owing to the circular boundary (cf. Vol. II, p. 244, and p. 13 of the present volume) are greater the smaller the raindrops are.

If we are to build up the rainbow formed by the white light of the sun from its elements on the basis of the above considerations, we must construct the diffraction pattern of fig. 6 for each colour and for each size of drop and then superpose the individual images in such a way that the point A of each image is associated with the maximum angle of deviation corresponding to its colour; thus, e.g., for red light the point A must lie at 42° 30′, for blue light at 41°. The result will be an extremely complicated mixture of colours, containing pure red only in the neighbourhood of the outer edge; inwards this gradually becomes mixed with other colours, the intensity of the red component diminishing at the same time. At the place where the first maximum for blue occurs the mixture of colours approaches blue, but the pure blue of the spectrum does *not* occur. At a greater distance from the edge the mixture of colours becomes still more irregular; the succession of colours in the supernumerary bows varies very much, until finally a complete mixture of all the components gives white.

As the maximum angle of deviation is constant for each colour, the outer edge of the primary bow always appears at the same angle; as, however, the distances of the intensity maxima for each separate colour depend on the size of the drops, the appearance and succession of colours vary from one rainbow to the next, especially as regards the supernumerary bows.

If the diameter of the drops is less than 0·1 mm., the individual diffraction maxima move so far apart that they overlap to a great extent for all colours even in the first maximum, although the diffraction pattern begins at a different point (A in fig. 6) for each colour; the bands of colour become broader and duller simultaneously. The very minute droplets in mist may give rise to a white rainbow instead of the ordinary coloured rainbow.

The **secondary bow** is due to that part of the pencil of rays which is again reflected at C (fig. 5), i.e. is reflected twice in all. This pencil of rays mostly emerges at E in the direction EF and meets the screen at F. If the "drop" is raised or lowered the spot always remains at some distance from S. It reaches the least distance from S when the angle between EF and the original direction of the ray is the supplement of 51°, the corresponding angle of incidence being 72°; as in the primary bow, the dispersion phenomena are accompanied by diffraction phenomena.

Fig. 7 shows how all the raindrops in a cloud contribute to the combined phenomenon visible to the individual observer. Only a limited pencil of rays from each individual drop can reach the eye

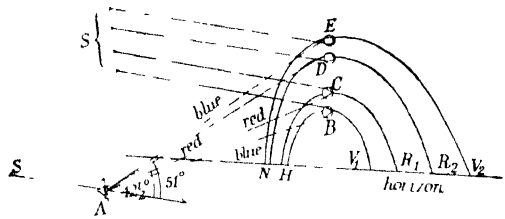


Fig. 7 Formation of the rainbow

A; each individual drop appears only in a colour which depends on the angle which the ray from the drop to the eye makes with the line from the eye to the counter-sun, but is, of course, surrounded by the diffraction phenomenon discussed above (very faint for a single drop), which depends on the size of the drop. Drops for which the angle just referred to is 42° 30′ appear red and form the outer edge of the primary bow. Next in inward order come the drops coloured green and blue. In the same way the drops for which the angle is 51° also appear red; they form the inner edge of the secondary bow, and are followed outwards by drops of the other colours. Between the primary bow and the secondary bow there is a comparatively dark region. Since the interference phenomena depend on the size of the drops, the details of the appearance of the rainbow (succession of colours, depth of tints), as mentioned above, vary with the size of the drops.

CHAPTER XII

Physiological Optics

1. The Visual Process.

When light falls on the retina, a chemical change takes place in its tissues, which is manifested by a change of colour in the retinal pigment or **visual purple.** How this chemical change gives rise to the sensation of light is not known. Immediately after being affected by light, the retina loses its power of vision at the region excited, owing to the chemical change in the visual purple. This is very well shown by the fact that we are dazzled on looking at an intensely bright light. The chemical change is reversed by the circulation of the blood. If it has been so marked that the normal state is not restored for some time, we experience particularly clear after-images, which disappear only by degrees.

Positive After-images, however, occur quite normally after every light-stimulus and are usually of short duration. A glowing piece of coal swung round in a circle no longer appears as a point but as a line, and when rapidly moved as a complete circle of light. The sensation of light does not cease immediately the light-stimulus does, but continues for a finite length of time after it. For the same reason, if pictures are rapidly altered they are no longer perceived separately, but one mingles with the next.

Use is made of this fact in the **cinematograph,*** in which a series of separate pictures of a certain action are brought in front of the eye at a definite rate (of about 16 pictures per second). Thanks to the existence of positive after-images, this series of pictures is experienced as an unbroken series of impressions, i.e. as a record of the successive motions. The rate at which the pictures are made to succeed one another is so chosen that the last vestige of the positive after-image of one picture has not entirely vanished when the next picture is presented to the eye.

* Gr., *kinema*, motion. The cinematograph is a development of stroboscopic illumination (Vol. II, p. 236), which goes back to PLATEAU (and to STAMPFER at about the same time (1823)). The first noteworthy attempt at producing moving pictures was the subjectively-used stroboscope of ANSCHUTZ (1845–1907). [The commercial history of the moving picture began with EDISON's kinetoscope and the invention of the Kodak celluloid film by EASTMAN.]

Negative After-images occur when a light-stimulus acts for some time on the same part of the retina, i.e. when the eye looks straight at a brightly illuminated object for some time and then at a surface which is feebly but uniformly illuminated. We then see an after-image in which light and dark are reversed. As the after-image of the glowing filament of an electric lamp we see a dark filament on a bright ground; as the after-image of the dark cross formed by the bars of a window against the bright panes we see a bright cross and dark panes.

The existence of after-images is satisfactorily explained by the assumption that a chemical change occurs at the back of the eye. In the case of the positive after-image the chemical change has been so marked that it cannot be reversed by the action of the circulation immediately the light-stimulus ceases. In the case of the negative after-image the affected parts of the retina become fatigued, and as a result the unaffected parts are more sensitive to light than those affected by prolonged exposure to light. The fatigue is not completely dissipated by the circulation for some little time, possibly even minutes.

2. Colour Vision.

Three-colour Theory.—A number of theories to explain *colour vision* have been put forward, but no one of them seems to explain all the phenomena completely. In the *Young-Helmholtz three-colour*

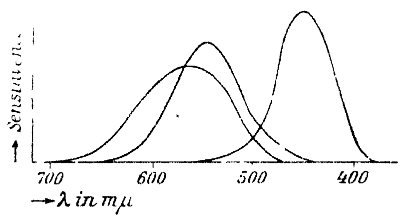


Fig. 1 —The variations of sensitiveness to the three colour components

theory, put forward originally by THOMAS YOUNG in 1807, redeveloped by HELMHOLTZ in 1867, and subsequently extended by other workers, it is assumed that the retina contains, in addition to **rods** which are *insensitive* to *colours* and *sensitive* only to *differences in brightness*, three different kinds of nerve endings (**cones**; see p. 113) *sensitive* to *colours*. One kind of cone is sensitive to reddish-purple, another to green, and the third to blue bordering on violet (fig. 1). According to J. VON KRIES * (1894) the yellow spot contains cones only, whereas rods predominate in the marginal parts of the retina.

If a ray of light of a single spectral colour falls on the retina, it excites all three kinds of cones, but in different degrees according to its colour. A red ray chiefly excites the cones sensitive to red, the two other kinds being excited only

* Professor of Physiological Optics at Freiburg im Breisgau.

feebly. According as the red has a tinge of yellow or blue, the cones sensitive to green or to blue are excited also. Yellow light makes itself manifest by exciting the cones sensitive to green and the cones sensitive to red to about the same extent, but the cones sensitive to blue only slightly. The sensation of white arises when all three kinds of cones are excited to the same extent.

The fact that two *complementary colours* (e.g. blue and yellow) together give white is at once intelligible on this theory.

In agreement with the Young-Helmholtz theory there is the fact that in making coloured illustrations it is possible to produce any shade of colour from three components (**three-colour printing**). The three colours used are red, yellow, and blue. The original in its natural colours is photographically split up into the three coloured components by the use of panchromatic plates and light filters which let through only a definite part of the spectrum. Then three different half-tone plates are produced, corresponding to the three photographs, and are used to print the three colours yellow, red, and blue on top of one another. The spectra in the coloured plate (facing p. 158) are reproduced by this process.

Three-colour Photography by the Maxwell-Ives Process.—Coloured pictures may be thrown on a screen by a similar process. The object is photographed three times through red, green, and blue glasses respectively. If lantern slides are made from the three photographs and are projected separately by different lenses on to the same part of a screen, the same glass filter being placed in front of each lantern objective as was used in making the photograph, the superposition of the three coloured images reproduces the variegated colours of the original object very well. In this projection the colour yellow is produced (additively) by means of a mixture of red and green light; this is surprising to painters who are used to working with pigmentary colours, as a red pigment and a green pigment do *not* give yellow (see p. 156).

Autochrome Photography.—Glass plates are coated (e.g. by the process due to the LUMIÈRE brothers, Lyons) with a very thin layer containing starch granules coloured red, green, and bluish-violet and only about $\frac{1}{100}$ mm. in diameter, the coating being so thin as to consist of only a single layer of granules and so uniform that at almost every point there are three differently-coloured granules adjacent. The three colours are so selected that their additive mixture (Chap. VII, § 2, p. 157) is a fairly pure white; the granules are so small and distributed so closely and uniformly that the coated plate appears white to the eye. A very thin layer of photographic emulsion (silver bromide) is then poured on top of this layer. Contrary to the usual practice in photography, the object is photographed through the glass. A yellow screen is used, in order that the blue light with its stronger photographic action may not have too strong an effect. Now if red light falls on a small area of the plate, say one square millimetre, it will pass through the red granules but will be absorbed by the neighbouring granules of the other colours. Yellow light is about half transmitted by the green granules and half by the red, but not by the blue; similarly for other colours. Only the rays transmitted by the granules can affect the sensitive photographic layer (which is panchromatic, or equally sensitive to all colours (Vol. V)). When the plate is developed by the usual method it will accordingly show blackening behind a red granule if the granule was illuminated by red light; less marked blackening is produced behind the red and green granules when yellow light falls on them, and so on. If the plate is then fixed, the red granule becomes opaque to light, while the neighbouring blue and green granules are transparent, as in their case the silver granule is wanting. Hence when the plate has been developed and fixed in the usual way, bluish-green light will be transmitted at points which were illuminated by red light, and so on. Thus the picture obtained when looked through exhibits not the right colours, but the colours complementary to them. To obtain the correct colours it is necessary to reverse light and dark.

After development the plate, without being fixed, is placed in a solution which dissolves the silver bromide reduced in the development but not that which is still unaffected by light. The plate is then exposed to a strong light and redeveloped. Silver granules are now formed in front of the starch granules which let through no light when the plate was originally exposed in the camera. A positive is thus obtained, which when looked through exhibits the correct colours. These Lumière transparencies are quite good, especially when projected on a screen; it is, however, impossible to make prints from them in the ordinary way. If they are looked at under the microscope, or if they are projected on a screen and examined at close quarters, the starch grains of various colours can be seen separately. The process does not give a good reproduction of spectra; the spectrum is split up into three parts separated by dark bands corresponding to the colours of the starch granules. Moreover, certain regions of the spectrum are absorbed by all the granules. As an example we may mention sodium light; a bunsen flame coloured with a sodium salt does not come out yellow in the photograph.

In the Agfa colour plates the fine particles which act as a light filter consist of an emulsion of coloured droplets of resin. As compared with the starch screen these plates have the advantage of greater transparency. Processes have recently been used, especially in the production of cinema films, in which only *two* complementary colours are employed; the results are surprisingly good.

Colour-blindness.—Deviations from normal colour vision are referred to as *colour-blindness*, the most common type of which is *red-green blindness* (Daltonism; see Vol. II, p. 19). It is manifested by the fact that the person afflicted with it cannot distinguish between red and green; for example, a ripe strawberry appears to him to have the same colour as the green strawberry leaves. In another type of colour-blindness, which is much less common, the person is unable to distinguish between blue and yellow. Total colour-blindness, in which no colours are perceived at all, is very rare.

Disturbances of colour-vision are very common and much commoner among men than among women (who on the average have greater visual acuity in other respects also). It is said that in Germany 4 per cent of the men are markedly colour-blind, a further 6 per cent have a noticeable colour defect, and 10 per cent a defect which can be detected only by careful testing, so that 20 per cent have abnormal colour vision. Our knowledge of the sensations of the colour-blind has been notably advanced by the rare cases in which one eye is colour-blind and the other normal; for it is only in such circumstances that an exact description can be obtained of the nature of the deviations from normal sensation. It may be mentioned, however, that the outer regions of the normal eye are incapable of perceiving colour. If a person under experiment is made to look intently and directly at a point without moving the eyes, and coloured discs are gradually brought farther and farther into his field of view, these discs are seen as such long before the person can tell their colour (or shape). In this way it is also easy to establish the fact that the regions of the retina which are capable of perceiving the various colours differ in size. Certain disturbances of colour vision may also arise from the use of drugs.

It seems to be established that persons with red-green blindness see all the longer-waved colours, i.e. red, orange, yellow, and yellowish-green, as *yellow*, the shorter-waved colours, i.e. greenish-blue and bluish-violet, as *blue*, and the intermediate zone, green to greenish-blue, as neutral *grey*.

Colour-blindness is explained by the assumption that all the types of colour-

sensitive cones are not equally well developed in the retina. **Lummer** (1904) explains total colour-blindness by the assumption that in such cases the yellow spot contains rods (see above), whereas in persons with normal colour vision it contains cones only.

Vision by the Rods.—The rods, although incapable of detecting colour, are more sensitive to light than the cones. Hence when the illumination is feeble, e.g. in a landscape seen by moonlight, vision takes place entirely by means of the rods, and hence no colours are seen; all objects appear grey, differing only in intensity. If the eye is directed straight at an object, its image is formed on the yellow spot, which is devoid of rods. For this reason objects looked at directly in a poor light may vanish, e.g. a very feeble star disappears when one looks straight at it, but reappears when one looks at a point near it.

As a result vision in a poor light is of an uncertain character. A nervous and superstitious person may imagine he sees ghosts at night which slip away and vanish when he tries to observe them carefully. The phenomena of the so-called *grey glow* also depends on vision by the rods. If a wire is heated electrically and watched by an observer in a perfectly dark room, before reaching dull red heat it becomes visible with a peculiar light, which excites the rods before the less sensitive cones react at a slightly higher temperature (see Vol. V

Colour and Intensity.—For the intensity of the colours of the spectrum see fig. 19, p. 31. When colours gradually fade as the sun sets the sky remains distinctly blue long after other colours are indistinguishable. At other times the sensitiveness to blue also persists longer as the intensity diminishes than that to other colours (Purkinje's phenomenon). As the intensity increases all colour sensations eventually pass into the sensation of white. The tint of the colour seen therefore depends on the intensity of the stimulus.

Coloured After-images.—The occurrence of *coloured after-images* (§ 1, p. 269) may also be explained by the Young-Helmholtz theory. If, for example, we look steadily at a brightly coloured object, e.g. a strip of red paper, lying on a uniform grey background, for about 10 to 20 sec. without moving the eye from the place, and the red paper is then removed suddenly, an after-image consisting of a bluish-green strip of the same shape is seen at the place where the paper was. Owing to fatiguing of the cones sensitive to red, the sensation of red derived from the uniform grey illumination is enfeebled and the subjectively perceived bluish-green arises from the green and blue components, which are stronger (because the cones on which they fall are not fatigued).

3. Optical Illusions.

Contrast effects, as a result of which an object invariably appears darker on a bright ground than it does on a dark ground, are partly due to physio-

logical causes. If we look at a bright surface, the *whole* of the retina is subject to a definite amount of fatigue which does not occur if we look at a dark surface. The whole retina is therefore more sensitive when a dark surface is looked at than when a bright surface is looked at: hence an object lying on a dark surface is perceived by a less fatigued retina, i.e. it appears lighter. The contrast effect may be well observed by laying a strip of grey paper on a black surface and a similar strip on a white surface; the former strip will appear lighter than the latter.

If we look at the outline of a wood which stands out dark against the sky and then raise the eyes a little higher, we see a bright rim parallel to the outline of the wood, which appears brighter than the rest of the sky, because the light from it is falling on a rested or less fatigued portion of the retina. Owing to continual slight involuntary rotations of the eyeball, a dark object usually appears surrounded by a brighter rim and, for the same reason, a bright object by a darker

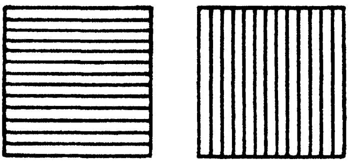


Fig. 2.—Illusion of size

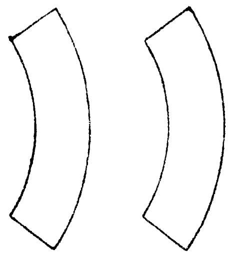


Fig. 3.—Illusion of size

rim. A painter can greatly increase the naturalness of his pictures by taking this phenomenon into account.

Colour contrasts are also partly due to the same cause. A piece of grey paper lying on a red surface appears greenish, whereas the same paper on a green background looks reddish. In the former case the retina has been fatigued all over for red light, in the latter case for green light.

Vividly-coloured objects appear surrounded by zones of the complementary colour, analogous to the bright and dark margins mentioned above.

The Colours of Shadows.—If we illuminate a white surface with red light, the shadow of a stick placed in front of the surface appears greenish. The sky behind red clouds at sunset appears green. The grey night sky often appears of a deep blue when looked at through the window of a room illuminated by a yellow paraffin lamp or electric light.

In all contrast effects there are psychological causes at work also; for when slight differences are directly compared they are apt to be overestimated.

Illusions of magnitude are also of interest; these arise from purely psychological causes even when the two objects compared are identical and are looked at from the same distance (i.e. have the same apparent magnitude). Thus of the two squares of fig. 2, which are shaded in different ways, the left-hand one appears higher than it is broad, whereas the right-hand one appears broader than it is high; similarly, the left-hand sector of a circular ring in fig. 3 appears larger than the right-hand one, although the two are congruent. The fact that we may be subject to illusions of shape also is shown by fig. 4.

Irradiation.—Bright objects, particularly when the eye is not focused exactly for them, appear larger than dark objects of the same size, because each luminous point gives rise to a circle of confusion on the retina, which apparently magnifies the bright boundary. The bright crescent of the new moon appears to

belong to a larger circle than the rest of the surface of the moon when the latter is visible at the same time. The brightly luminous filament of an electric lamp appears several times as thick as it really is. A peculiar illusion arises chiefly in connexion with the apparent magnitude of the sun. In all pictures in which the sun appears it is found that the artist, following his sensations, has painted it much too large relative to other objects shown in the picture; often it is as much as forty times too large. In photographs of landscapes which include the sun, the latter always appears "unnaturally" small, as the lens of the camera depicts it on the correct scale.

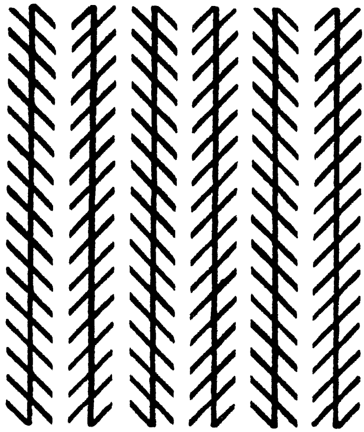


Fig. 4 --Another optical illusion

Judgment of Distance.—This is subject to many illusions, most of which have a psychological basis. In a hazy atmosphere one is apt to overestimate distances owing to the blurring of the outlines of objects, whereas in the clear air of high mountainous regions, and also at the seaside in clear weather, objects appear nearer than they really are. The power of estimating distances accurately and reliably can only be acquired by systematic practice.

That the sun and the moon appear larger near the horizon than at the zenith is due to the fact that the sky appears to us not as a hemisphere but as a flattened segment of a sphere (fig. 5). Two reasons may be given for this: (1) distant objects on the earth's surface always appear somewhat indistinct, as the rays from them have to travel over long distances in the hazy layers of air next the ground; hence they seem more distant than they really are (atmospheric illusion). On the other hand, the more or less vertical rays traverse only a thin layer of the hazy part of the atmosphere; stars near the zenith appear sharp and relatively nearer. (2) Angles of elevation are overestimated, perhaps because the muscular effort required to move the axis of the eyes in a vertical plane is greater than for a horizontal plane. For example, if we try by ocular estimation to fix a point in the sky which is equidistant from the zenith and the horizon, i.e. which has an angular elevation of 45°, we place it much too low, at about 22° (as we may easily convince ourselves by subsequent measurement). Now as the distances of points in the sky near the horizon appear greater than in the neighbourhood of the zenith, objects which actually subtend equal angles at the eye appear larger the nearer they are to the horizon.

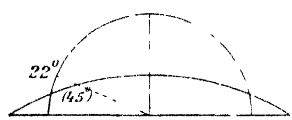


Fig. 5 --Apparent shape of the sky

The fact that the eye has to accommodate to a definite extent to see objects at a definite distance sharply has only a minor bearing on the judging of distances. Our sensation of muscular effort in the ciliary muscle which brings about accommodation only affects our judgment of distances at close quarters (within one metre). Greater importance attaches to the angle included between the axes of the two eyes when we look at an object with the two eyes simultaneously or observe it successively from two different points at equal distances from the object (§ 4). Distant objects are then apparently displaced less than nearer ones (parallax; see Vol. I, p. 10).

4. Binocular Vision.

In binocular vision a retinal image is formed in each eye; under normal conditions, however, the two retinal images are consciously perceived as a *single* impression, especially if the images are formed on the yellow spots of the two eyes. This is explained by the fact that the two optic nerves intersect behind the eyes; each nerve-ending in one eye has associated with it a nerve fibre which coalesces with the nerve fibre of the corresponding nerve-ending in the other eye. Two points of the retina such that the corresponding nerve fibres coalesce with one another are called **corresponding points** or **conjugate points** of the two retinæ. When we look at an object M (fig. 6), we involuntarily direct the eyeballs in such a way that the point of the object that we are fixing the eyes on gives rise to images at the yellow spots of the two eyes (g_l, g_r), that is,

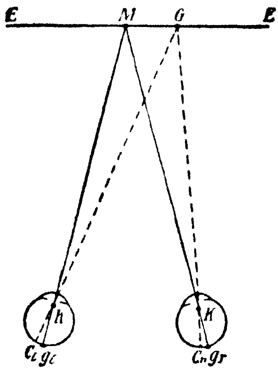


Fig. 6.—Production of the conjugate images in the two eyes

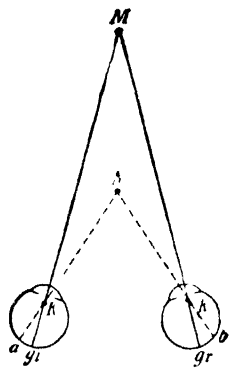


Fig. 7.—An object N lying nearer than the fixed object M appears double.

so that the point looked at lies on both optic axes. A point G on the plane EE which is parallel to the line joining the two eyes, i.e. has all its points at approximately the same distance from the two eyes, like M, gives rise to an image c_l in the left eye and an image c_r in the right eye. These two points are conjugate points of the retinæ of the two eyes, for the two-fold image of G produces only a single sense impression. In both eyes the conjugate points lie on the same side, both to the left or both to the right of the yellow spot. Similarly all images formed on the retina by other points of the plane EE are produced at conjugate points in the two eyes.

This is not true if the objects are situated like M and N in fig. 7. If the object M is looked at fixedly with the two eyes, the nearer point N gives rise to two retinal images a and b which are on *different* sides of the yellow spot in each eye. These points are not conjugate points. Hence if we look fixedly at a distant point with the two eyes, we see a nearer object double. Similarly if we fix our gaze on a near object a remoter object appears double.

If the distance in depth (i.e. along the line of sight) between M and N is not too great, the two different retinal images of the point N coalesce when we look at M, and form a single blurred image.

From the imperfection of the superposition of the two images a and b we infer that the distance of the point N from the eye differs from that of the object M which we are looking at. On this depends our capacity for directly perceiving

differences in depth, i.e. of seeing natural objects solid (*stereoscopic* * *vision*).

Stereoscopic Vision. The retinal images of an object with "depth" which

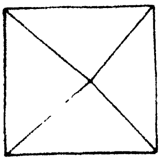


Fig. 8.—The two images of a pyramid

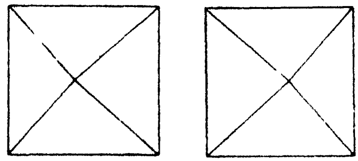


Fig. 9.—The two images of a hollow pyramid

are produced simultaneously in the two eyes are not the same. For example, if we look from above at a square pyramid standing with its square base on a table, the retinal image in the left eye resembles the left-hand side of fig. 8, that in the right eye resembles the right-hand side of the same figure. The left eye sees more of the left-hand side of the pyramid and the right eye more of the right-hand side; in both cases the vertex of the pyramid appears displaced towards the nose side. Conversely, when we look into a hollow pyramid the retinal images in the left and right eyes respectively are as shown in fig. 9; the vertex of the pyramid appears displaced away from the nose.

If we look with the left eye at the left picture and simultaneously with the right eye at the right picture in fig. 8 or fig. 9 (aiding the vision by placing a piece of paper at right angles to the line joining the two pictures and looking on either side of it) so that the two component pictures fall on conjugate points of the two retinæ, we obtain the same total impression as if we had a three-dimensional model of a raised or hollow pyramid in front of us.

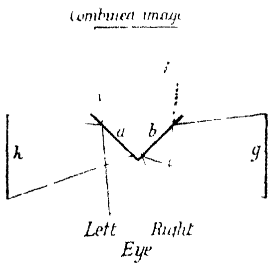


Fig. 10.—Path of the rays in Wheatstone's mirror stereoscope

The Stereoscope.—To facilitate the observation of the two pictures separately, one by each eye, various types of special apparatus have been designed.

In Wheatstone's † mirror stereoscope (1838) the two component pictures are set up at k and g and reflected by the two mirrors a and b at right angles to one another. In the mirror a the left eye sees only the picture at k, while in the mirror b the right eye sees only the picture at g. The combination of the two component pictures to form a single picture is explained by the paths of the rays in fig. 10.

In Brewster's prism stereoscope (1849), shown in fig. 11, the action of the two prisms P_1 and P_2 is such that the two eyes l, r see the component pictures L and R as if they were superposed. The usually accepted view is that these images are superposed in a plane D *behind* the image plane; a different view, however, has been put forward by French,‡ who suggests that the combined image is formed *in* the image plane, just as in the mirror stereoscope already described. The prisms are obtained by cutting a large convex lens in half and inserting the parts into the stereoscope in such a way that the sharp edges face one another.

* Gr., *stereos*, a solid body; *skopein*, to see.

† Charles Wheatstone (1802–75), originally an instrument-maker, was made professor at King's College, London, as a result of his numerous discoveries and inventions, and subsequently lived as a private gentleman on the income derived from his inventions; he was one of the most successful of practical physicists.

‡ *Transactions of the Optical Society*, Vol. XXIV, p. 226 (1922).

Thus not only are the rays given the desired path but the perspective (p. 125) of the pictures observed is improved and they are magnified.

Stereoscopic pictures may also be recombined by using an opera-glass from which the eyepiece has been removed, if the objective is held immediately in front of the eyes, i.e. if one looks through the "wrong end" of the opera-glass.

Helmholtz's telestereoscope, of which a good idea may be got by imagining k and g in fig. 10 replaced by two mirrors parallel to a and b, has the effect of considerably magnifying the distance between the observer's eyes; prism binoculars act in a similar way (p. 135). If one looks at a landscape through these instruments its depth appears greatly extended. A further effect of the magnification of the distance between the eyes, probably psychological, is that the apparently distant parts of the landscape then appear smaller than when they are observed directly, although the angles subtended at the two eyes are not diminished.

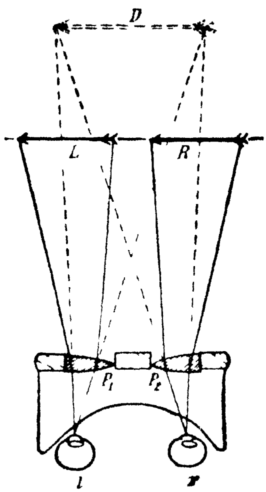


Fig. 11.—Brewster's prism stereoscope

The Stereogoniometer. In aerial survey work, the methods of mapping from air photographs suffer in most cases from two limitations: they require a fairly dense ground control of heights for contouring, and they cannot be applied to the survey of hilly country, unless accuracy be sacrificed, without certain mathematical complications. These limitations are overcome in the stereogoniometer designed by Barr and Stroud Ltd. and embodying the invention of Dr. Henry Fourcade. This instrument has been used in the production of topographic maps by the Geographical Section of the British War Office.

Essentially, the instrument consists of a three-dimensional measuring stereoscope, which may be used to establish scale and level on a bare minimum of ground control, a process comparable with the use of a theodolite in executing secondary triangulation on the ground.

If two photographs are taken without altering the principal point or principal distance of the camera, there are five ways in which correspondence between any common points may be destroyed:

(1) and (2) *Either* photograph may be rotated in its own plane about an axis coinciding with the plate perpendicular;

(3) *One* photograph may be rotated about an axis coinciding with its base line (or line of flight);

(4) and (5) *Either* photograph may be rotated about an axis passing through its principal point and perpendicular to the base line.

The principle of the stereogoniometer is based on the fact that geometrical correspondence can be restored by means of these same five movements and, once correspondence has been restored, the two photographs will occupy the same relative position to each other and to the base line as they occupied at exposure

In the instrument, the counterpart of the base line is known as the *polar axis*, and the axes mentioned in (4) and (5) are known as *declination axes*.

The essential parts of the instrument are shown in fig. 12.

The goniometers (9), in which the air photographs are placed, are carried on declination axes at right angles to, and carried by, the polar axes. They are provided with camera lenses and are generally similar to skeleton cameras.

Setting movements in the instrument are as follows.

(*a*) Rotation of either of the rings (10), carrying the photographic plates, about the plate perpendicular. This rotation can be finely adjusted and measured.
(*b*) Rotation of either goniometer about its declination axis.
(*c*) Polar rotation of the right-hand goniometer only, about its polar axis, performed by means of the head (12).

The photographs, illuminated by artificial light, are examined stereoscopically in a prismatic telescope. Beams of light from the photographs are refracted by the goniometer lenses (8) and, emerging as approximately parallel beams, are

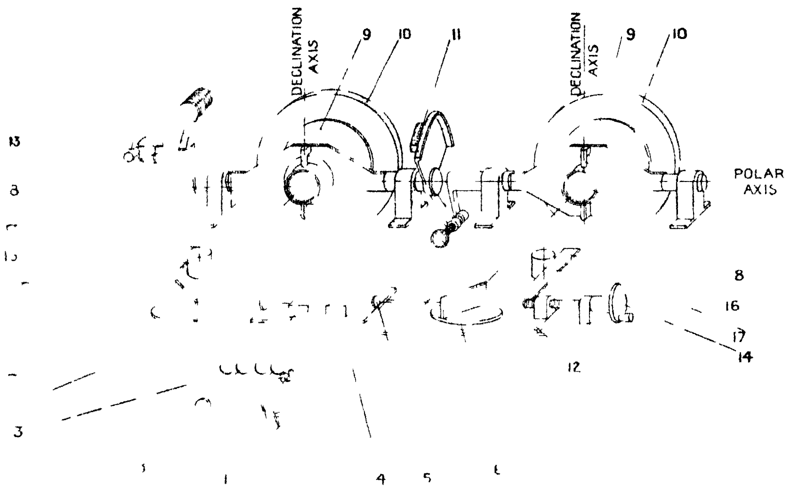


Fig. 12.—Diagram [illegible] the stereogoniometer (Barr & Stroud Ltd.)

reflected by the plane mirrors (6) and (7) into the objectives (5) of a binocular telescope provided with floating marks (3) in the focal planes of the objectives. Before forming images in the plane of (3) the light is reflected by right angled prisms (4). Images formed in the plane of (3) are examined, together with the floating marks, through the eyepiece lenses (1). Two types of interchangeable marks (3) are provided. The lenses (2) are erecting lenses.

To enable different areas of the photographs to be brought under examination the following additional *observational movements* are provided.

(*a*) The handwheel (13) rotates both goniometers simultaneously about the polar axis, changes in "polar bearing" being read by means of the scale and vernier (11).
(*b*) The handwheel (15) translates the two mirrors (6) and (7) together in a left and right direction. The handwheel (14) similarly translates the right hand mirror (6) only. As each mirror carrier is translated the mirrors are automatically rotated by the amount necessary to reflect the fixed lines of collimation of the telescope into the goniometer lenses.

Twice the amount of rotation of the mirrors is equal to the corresponding

rotation of the sighting rays between the lenses (8) and the mirrors and is read directly—without multiplication—on declination scales on the mirror carriers.

When the instrument is properly adjusted, a pair of photographic images will be in correspondence when they appear to lie an equal distance above or below the floating marks.

By working the five setting movements, the whole of an overlap formed by two photographs can be brought into correspondence.

Stereoscopic eyepieces in microscopes are frequently used to examine the spatial configuration of the object under investigation. This is possible in spite of the extremely small range in depth of the microscope, as owing to the high numerical aperture of the objective the angle between the two directions from which the preparation may be looked at by the two eyes can be made quite large. The subdivision of the pencil of rays into portions traversing different paths may be brought about in a number of ways which we cannot discuss in detail here.

TABLE I.—REFRACTIVE INDICES. VALUES OF THE SPECIFIC DISPERSION $\theta = \mu_H - \mu_C$

Fraunhofer Line	A	B	C	D	E	F	G	H	θ
Wave-length in thousandths of a micron	760	686·7	656·3	589·0	527	486·1	430·8	396·8	—
Water	1·329	1·331	1·332	1·334	1·336	1·338	1·351	1·344	0·012
Alcohol	1·359	1·360	1·361	1·363	1·365	1·367	1·371	1·374	0·013
Benzene	1·493	1·495	1·497	1·503	1·507	1·514	1·524	1·536	0·039
Ether	1·351	1·355	1·356	1·358	1·360	1·362	1·366	1·371	0·014
Oil of cassia	1·586	1·592	1·596	1·605	1·619	1·634	1·665	1·700	0·105
Carbon disulphide	1·610	1·616	1·620	1·629	1·642	1·654	1·679	1·702	0·082
Light crown glass	1·510	1·512	1·513	1·515	1·519	1·521	1·527	1·531	0·018
Heavy flint glass	1·735	1·741	1·743	1·752	1·762	1·772	1·792	1·811	0·068
Iceland spar { ordinary ray	1·650	1·653	1·655	1·659	1·664	1·668	1·676	1·683	0·028
Iceland spar { extraordinary ray ..	1·483	1·484	1·485	1·486	1·489	1·491	1·495	1·498	0·013
Quartz { ordinary ray	1·539	1·541	1·542	1·544	1·547	1·550	1·554	1·558	0·016
Quartz { extraordinary ray .. .	1·548	1·550	1·551	1·553	1·556	1·559	1·564	1·568	0·017

Table II.—Rotatory Power of Quartz

Fraunhofer Line	A	B	C	D	E	F	G	H
Rotatory power of quartz	12·7°	15·7°	17·3°	21·7°	27·5°	32·7°	42·6°	51·2°

A solution of cane sugar containing x gm. sugar in 100 c.c. of solution, when in a layer l cm long, rotates the plane of polarization of sodium light through an angle α given by $0{\cdot}0665\ xl$

Table III.—Atmospheric Refraction

Apparent elevation	0°	1°	2	3	4°	5	6°	8°	10°
Atmospheric refraction	34′ 54″	24′ 25″	18′ 9″	14′ 15″	11′ 39″	9′ 47″	8′ 23″	6′ 30″	5′ 16″
Apparent elevation	12°	16°	20°	25°	30°	40°	50°	60°	80°
Atmospheric refraction	4′ 25″	3′ 32″	2′ 37″	2′ 3″	1′ 40″	1′ 9″	48″	33″	10″

EXAMPLES

Examples A (pp. 1–79)

1. How is the rectilinear propagation of light explained by the wave theory? Discuss the diffraction effects produced by (*a*) a straight edge, and (*b*) a narrow object such as a wire.

(*C. Tripos, Pt. I.*)

2. Give the elementary theory of diffraction at a straight edge, and show how you would demonstrate the phenomenon experimentally. Explain how measurements of the diffraction pattern observed in the shadow of a wire can be used to form an estimate of the wave-length of light.

(*Ox. F.*)

3. Give an account of Huygens' theory of wave propagation and interference.

(*Br. F.*)

4. Show how the nature of the diffraction fringes seen in the shadow of a straight edge can be explained.

Light issuing from a slit throws a shadow of a needle, parallel to it, on to an eye-piece. Describe what is seen in monochromatic light, and discuss the effect of diminishing the breadth of the needle.

(*Br. F.*)

5. What is the international standard of illuminating power? A lamp of 100 candle-power, all the light from which can be considered to come from a point, is placed at the centre of a sphere of radius 12 cm. The inside of the sphere is blackened and a circular hole 1 cm. in diameter is cut in it. What is the intensity of illumination on a screen placed 48 cm. from the lamp and perpendicular to the radius passing through the centre of the hole? To what value is this changed if a convex lens of focal length 10 cm. is placed over the hole? (See Chapter IV.)

(*Sheff. Inter.*)

6. Describe a form of photometer for the accurate comparison of the illuminating power of two sources of light. How would you use it to determine the mean horizontal candle-power of an incandescent lamp?

(*Lond. Inter.*)

7. Distinguish between illuminating power and intensity of illumination. A 100 candle-power lamp is suspended 20 ft. above the middle of a road which is 20 ft. wide and horizontal. Find the intensity of illumination at a point on the edge of the road nearest the lamp.

(*Lond. Inter.*)

8. To a man looking into still water the depth appears less than it really is. How does this illusion arise? What would be the apparent depth of a pond 8 ft. deep, the refractive index of water being 1·33?

(*Br. Matr.*)

9. Explain why a pool of water appears less deep than it actually is, and deduce a relation between the index of refraction, the real depth and the apparent depth. Describe how this relation could be tested experimentally.

(*Br. Inter.*)

10. Describe the construction of a sextant, and explain how you would use the instrument to measure the altitude of the sun above the horizon.

(*Birm. Inter.*)

11. Describe how you would determine the refractive index water-air by the air-film method. Show that the result is independent of the material of the plates used to confine the air film.

(*Br. F.*)

12. (1) What are the conditions under which the total reflection of light takes place?

(2) How would you use this property in determining the refractive index of a liquid experimentally?

(3) What practical use is made of the total reflection of light by prisms?

(*Br. Inter.*)

13. Describe how you would find the focal length of a convex mirror. A circular source of light 1 cm. in diameter is placed at the focus of a concave mirror of 30 cm. focal length. At what distance would the projected beam of light just be able to cover an aeroplane with a wing span of 60 ft.?

(*Birm. Inter.*)

14. What is meant by the focal length of a concave lens? How would you measure it experimentally?

A convex lens of focal length 6 cm. is placed in contact with a concave lens of focal length 18 cm. An object is placed 5 cm. from the combination. Where will the image be, and will it be real or virtual?

(*Lond. Inter.*)

15. Discuss the position, magnitude, and nature of the image of an object placed at distances varying from zero to infinity along the axis of (1) a concave spherical mirror, and (2) a convex spherical mirror.

A concave and a convex spherical mirror are placed with their mirror surfaces facing each other, at a distance of 10 cm. apart. The radius of curvature of each of the surfaces is 10 cm. An object 1 cm. long is placed perpendicularly to the axis of the mirrors and at a distance of 7 cm. from the concave mirror. Find the nature, position, and size of the image formed after two reflections.

(*Leeds Inter.*)

16. Distinguish between real and virtual images and explain the conditions under which a concave mirror gives rise to each. An object is 5 in. and the image 20 in. from a concave mirror. Find the radius of curvature of the mirror when the image is (*a*) real, and (*b*) virtual.

(*Br. Inter.*)

17. Show by a graph the relation between the distance of an object and that of its image produced by a thin convex lens. Include in your graph the relation both when the object is real and when it is virtual, and show by a diagram an arrangement for investigating the latter case experimentally.

Explain how you would determine experimentally the refractive index for sodium light of the material of a thin convergent lens.

(*Br. Inter.*)

18. The two perpendicular faces of a right-angled prism of glass are both 4 cm. square. Examine the effect of the prism on a narrow pencil of light incident normally on the middle of one of these faces from an illuminated pin-hole placed

24 cm. from the face, (*a*) when the face is truly plane, and (*b*) when it is convex with a radius of curvature of 120 cm. The refractive index of the glass is 1·5.
(*Sheff. Inter.*)

Examples B (pp. 79–149)

1. Prove that the focal length of a system of two lenses, of focal lengths f_1 and f_2, distant d apart, is F, where $\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}$.

A telephoto lens consists of a convex lens facing the object with a divergent lens behind it. The focal length of the first lens is 10 cm., of the second 4 cm., and they are 7 cm. apart. Find where the plate must be placed to photograph a distant object, and the size of the image if the object subtends 1°.
(*Man. Inter.*)

2. What are the cardinal points of a lens system?

A glass sphere of radius r and refractive index β is used as a thick lens, its aperture being made small by the use of suitable stops. Find the focal length and the position of the two principal points.
(*C. Tripos, Pt. I.*)

3. Define the cardinal points of a coaxial lens system, and explain how the position of the image of an object on the axis can be determined when the position of the cardinal points is known.

Deduce the positions of the cardinal points of a glass hemisphere in air.
(*Br. F.*)

4. What is meant by (1) the optical centre, and (2) the nodal points of a thick lens?

Deduce Helmholtz's formula which expresses the magnification in terms of the divergence of rays before and after refraction by a lens, and show that in the case of a thick glass lens in air the nodal points coincide with the points of unit magnification.
(*Br. F.*)

5. Explain the optical properties of the photographic camera.

When a camera is in focus for distant objects the ground-glass screen is 6 in. from the lens. When the bellows are fully extended the screen is 7 in. from the lens. What is the smallest distance from the lens at which an object can be photographed?
(*Br. Inter.*)

6. Describe briefly the optical lantern, explaining the function of its lenses.

A slide 3 in. square is to be projected on to a screen 100 ft. from the slide, the image to be 15 ft. square. Where must the objective be placed, and what must be its focal length?
(*Dubl. Inter.*)

7. Describe the human eye, and compare its action with that of a camera. What do you understand by long and short sight? How are these defects remedied?
(*Dur. Inter.*)

8. Deduce an approximate expression for the resolving power of a telescope. Why is it that stars which are not visible to the naked eye in broad daylight can be seen through a telescope?
(*Br. F.*)

9. What is meant by the resolving power of a telescope? Deduce an expression for it.

What effect has the eye-piece on the resolving power?
(*C. Tripos, Pt. I.*)

10. What is meant by the resolving power of a lens? If a lens of focal length f is used to form the image of a double star whose components subtend a small angle β at the earth, what must be the aperture (d) of the lens in order that the image may consist of two bright circles just touching one another?

(*Br. F.*)

11. Discuss the phenomenon of diffraction, illustrating your answer by considering in detail *one* of the following cases:

(a) The distribution of intensity along the axis of a circular aperture illuminated by monochromatic light from a point source on the axis.

(b) The distribution of intensity in the focal plane of a convergent lens illuminated by a parallel beam of monochromatic light, the aperture of the lens being limited by a slit.

(*C. Tripos, Pt. I.*)

12. Describe the construction of a modern field-glass, explaining how
(1) the production of colour effects by the object glass is avoided, and
(2) the erection of the final image is secured.

(*Br. Inter.*)

13. Survey the principles governing the resolving power of optical instruments. "The limit of resolution of the microscope is attained when the total magnification is about 900." Justify this statement.

(*Ox. F.*)

14. Draw a diagram showing the path of rays through a compound microscope when the image is formed coincident with the object at the distance of distinct vision.

If this distance is 28 cm., the distance of the object from the objective 4 cm., and the magnifying power of the instrument 14, find the focal lengths of the lenses.

(*Man. Inter.*)

15. Show that an object appears equally bright at all distances provided that it is an extended one. Show also that the image of such an object formed by a lens or a mirror cannot be brighter than the object itself.

(*Sheff. Inter.*)

16. Describe and explain the defects of vision known as long sight, short sight and astigmatism, and show how each of them can be corrected. A man whose "near point" is one metre requires spectacles to enable him to read small print at a distance of 20 cm. from his eyes. Calculate the nature and the focal length of the lenses required.

(*Birm. Inter.*)

17. Prove that in the case of refraction at a single spherical surface there is one position of the object point for which the image is also a point. State any application of this which is made in practice.

(*Lond. Hons.*)

Examples C (pp. 149–173)

1. Describe the construction of a spectroscope and show how you would use it to obtain a spectrum of sunlight.

How do you account for the fine dark lines which cross the spectrum of the light from the sun?

(*Birm. Inter.*)

2. What is the relationship between the angle of minimum deviation of light passing through a triangular prism of glass, the refractive index of the glass, and the refracting angle of the prism? How would you arrive at the equa-

tion connecting them? Explain the construction of a spectrometer. How would you use the instrument to measure the refractive index of a piece of glass in the form of a prism?

(*Birm. Inter.*)

3. Describe an experiment to show the composite nature of white light. Explain what primary and complementary colours are, and say how you would determine experimentally the colour which is complementary to green.

Why should yellow and blue pigments when mixed appear green while a mixture of yellow and blue lights appears white?

(*Sheff. Inter.*)

4. Describe the chromatic defects of a simple lens when used to produce an image of a source of white light.

How would you devise a convergent lens which is free from these defects?

(*Br. Inter.*)

5. Define dispersive power.

Given two lenses in contact, of focal lengths f_1 and f_2, and of materials with dispersive powers D_1 and D_2 respectively, find the condition for achromatism.

(*C. Tripos, Pt. I.*)

6. Deduce a formula for the variation with wave-length of the refractive index of a vapour in the region of an absorption band. Describe how the formula has been verified experimentally.

(*Br. Hons.*)

7. Describe and discuss the essential features of the quartz prism spectrograph. In such an instrument the quartz prism is equilateral, the sides being of length 7 cm., and the middle ray of the spectrum, $\lambda = 2963$, passes through the prism at minimum deviation. Determine the linear dispersion along the photographic plate and the maximum resolving power of this instrument in the region $\lambda = 2963$. The refractive indices of quartz for the wave-lengths $\lambda = 2746{\cdot}7$ and $\lambda = 3178{\cdot}8$ are $1{\cdot}5875$ and $1{\cdot}5729$ respectively; the focal length of the camera lens for $\lambda = 2963$ is 45 cm.; the surface of the photographic plate makes an angle of 22° with the axis of the lens. (See page 211.)

(*Sheff. Hons.*)

8. Describe and discuss the various defects of the simple lens.

(*C. Tripos, Pt. I.*)

Examples D (pp. 173–212)

1. What conditions are necessary for two beams of light to produce interference fringes? Explain how such fringes may be formed (*a*) with a plane mirror, and (*b*) with a biprism.

(*C. Tripos, Pt. I.*)

2. Discuss the production of interference fringes by two similar point sources of monochromatic light.

Explain carefully how you would obtain such fringes in practice, and how you would use them to measure the wave-length of the light employed.

(*C. Tripos, Pt. I.*)

3. Given two pieces of plane glass, mounted so that the angle and distance between their surfaces can be varied, describe how you would obtain (*a*) straight and (*b*) circular interference fringes.

Explain how you would measure the angle between the plates in the former case and the distance apart in the latter.

(*C. Tripos, Pt. I.*)

4. Describe the Michelson interferometer, and explain how it is used to obtain (*a*) straight fringes and (*b*) circular fringes with monochromatic light. Under what conditions may fringes be obtained with white light?

Indicate briefly any of the uses to which this instrument has been put.

(*Man. P.*)

5. Give the complete theory of the plane diffraction grating and sketch the form of the interference fringes in the case of a grating with a given small number of lines. What advantages result from the use of gratings ruled on curved surfaces? Describe a method of mounting such a grating.

(*Man. P.*)

6. Explain the action of the plane diffraction grating, and show how to calculate the angles at which the various spectra will occur. Distinguish between the dispersion and the resolving power of such a grating, and show what each depends on.

(*Man. P.*)

7. Describe a method of producing interference fringes with white light, indicating the special conditions required. What would be observed if horizontal fringes of this kind were viewed through a spectrograph with its slit (*a*) horizontal and (*b*) vertical?

(*C. Tripos, Pt. I.*)

8. Explain how the wave-length of the standard red cadmium line has been accurately determined.

(*Br. Hons.*)

9. Give the theory of (*a*) an echelon grating and (*b*) a Lummer-Gehrcke plate.

A certain echelon grating consists of 25 glass plates ($\mu = 1{\cdot}500$) of 20 mm thickness and step width 1 mm. Calculate the approximate order of the spectra observed and the resolving power of the grating for a wave-length $\lambda = 6 \times 10^{-5}$ cm., given that $\frac{d\mu}{d\lambda} = -400$, and assuming that the angle between the incident and emerging ray is small.

(*Br. F.*)

10. Describe a method of measuring the refractive index of a gas. In an experiment with a Jamin interferometer the length of the gas-filled tube was 20 cm. On changing the pressure of the gas by 70 cm. of mercury, 70 fringes passed the cross-wire of the observing instrument. Find the refractive index of the gas at a pressure of 76 cm. of mercury and at the temperature at which the experiment was carried out. ($\lambda = 5{\cdot}89 \times 10^{-5}$ cm.)

(*C. Tripos, Pt. I.*)

10. Explain the colour effects obtained from thin films of transparent substances under certain conditions. Indicate what these conditions are. What conditions are necessary before similar colour effects are obtainable with thick plates?

(*Liv. Hons.*)

12. An arrangement of equal parallel slits in a screen, spaced so that the width separating slits is double the width of a slit, is illuminated by light from a parallel, distant, slit. Find the value of the intensity of illumination transmitted by the set of slits in different directions.

(*Liv. Hons.*)

13. How is the "visibility" of a system of interference fringes defined? Describe Michelson's method of analysing spectral lines by measurements of the

visibility of fringe systems. Describe how observations of fringe visibility have yielded an estimate of the apparent diameter of stars.

(*Lond. Hons.*)

(If I_1 is the maximum intensity of light in a fringe system, and I_2 the minimum, the "visibility of the fringe system" is defined as the quantity $\frac{I_1 - I_2}{I_1 + I_2}$. If, with Michelson's apparatus, the spectral line has a fine structure, then as the path difference between the two interfering sources is increased the visibility of the fringe system changes periodically according to whether the fringe systems produced by the different components of the line are "in or out of step". See p. 196.)

14. Explain the respective effects of long path difference and semi silvering on the nature of fringes in the interference phenomena shown by two parallel glass plates.

(*Br. Hons.*)

15. Describe and give the theory of the formation of fringes obtained with a Fabry-Pérot interferometer. Discuss the advantages to be gained by half silvering the surfaces.

(*Lond. Hons.*)

Examples E (pp. 213–280)

1. Write a brief account of two methods of measuring the velocity of light, one an astronomical and the other a terrestrial method. Compare their advantages.

(*Sheff. Inter.*)

2. Describe how the velocity of light in air and in water has been measured directly and explain how the results support the wave theory of light.

(*Sheff. Inter.*)

3. State the principal facts of the polarization of light by reflection from a glass surface, and describe the experiments you would make to investigate them.

(*C. Tripos, Pt. I.*)

4. What is meant by (*a*) plane polarized light, (*b*) circularly polarized light, and (*c*) elliptically polarized light? How may each of these be produced? How would you distinguish between (*a*) unpolarized light and circularly polarized light, and (*b*) elliptically polarized light on the one hand and a mixture of plane-polarized light and unpolarized light on the other?

(*C. Tripos, Pt. I.*)

5. A thin wedge of very small angle, made of quartz, and cut so that the optic axis is parallel to the edge of the wedge, is viewed between crossed Nicol prisms in monochromatic light. Describe and explain the appearance of the wedge. How does the appearance alter as the wedge is rotated in its own plane between the Nicols?

(*Man. P.*)

6. Give a descriptive account of the optical properties of transparent crystals. Explain the interference phenomena which may be observed when crystalline plates are viewed in convergent polarized light.

(*Man. P.*)

7. How would you determine the specific rotary power of the plane of polarization with light of various wave-lengths for a specimen of quartz? How does the power vary with the wave-length? Explain Fresnel's theory of the pheno-

menon of rotation. What differences are there between the phenomena in quartz and those in glass in a magnetic field?

(*Lond. Hons.*)

8. Explain the formation of primary and secondary rainbows, and account for the orders of the colours.

Taking the refractive index of water to be 1·329 for red light and 1·343 for violet light, calculate the angular separation of the red and violet rays in the primary bow.

(*C. Tripos, Pt. I.*)

9. Lycopodium particles, which may be taken as spherical and of approximately the same diameter, are scattered on a sheet of glass, and a point source of monochromatic light is viewed through the glass. Discuss the formation of the corona which is observed.

What would be the effect of replacing the opaque particles by transparent discs all equal in area and thickness?

(*Ox. F.*)

10. Explain the occurrence of primary and secondary rainbows. Show that the region of the sky between them will appear completely dark, so far as the light coming from the raindrops is concerned. What are "supernumerary bows"?

(*Lond. Hons.*)

ANSWERS TO EXAMPLES

A **5.** (*a*) 434·1 lux. (*b*) 43,410 lux = 4·341 phot.

7. 0·179 foot-candles. **8.** 6 ft. **13.** 600 yards.

14. An erect virtual image 11·25 cm. from the combination.

15. An erect virtual image, 5 cm. high, and 15 cm. behind the convex mirror.

16. (*a*) 8 in. (*b*) 13·33 in.

18. The pinhole appears (*a*) 26·67 cm. and (*b*) 29·33 cm. behind the emergent face.

B **1.** 12 cm. behind the divergent lens; $\frac{2\pi}{9}$ cm.

2. The principal points coincide at the centre of the sphere; focal length equals $\frac{r\beta}{2(\beta-1)}$.

3. The lens is divergent with a focal length $\frac{r}{\mu-1}$; the principal points are (1) at the pole of the mirror, (2) at a distance $\frac{r}{\mu}$ from the plane surface inside the lens. The order of points is $F_1H_1H_2F_2$. **5.** 42 in.

6. The objective of focal length 19·36 in. must be placed 19·68 in. from the slide.

10. $d = 2{\cdot}44\ \lambda/\beta$. (Note that d is independent of the focal length f.)

14. Focal length of object lens equals $-3{\cdot}2$ cm.; of eyepiece $-11{\cdot}2$ cm.

16. Diverging lens of focal length 25 cm.

C **5.** $D_1/f_1 + D_2/f_2 = 0$.

7. Dispersion 15 Ångstrom units per mm.; resolving power 23,650. (Compare this with the data for a flint-glass prism given on page 212.)

D **9.** Order of spectrum 16,700; resolving power 420,000.

10. 1·000224.

E **8.** 1° 52′.

INDEX